LABOUR HETEROGENEITY IN A FARM HOUSEOLD MODEL
An Application to Nepalese Agriculture

Prem Jung Thapa

A thesis submitted for the degree of
Doctor of Philosophy
of
The Australian National University

June 1999

Dedicated to:

My parents,

and to Mac and Leo for getting me started on this journey

DECLARATION

The research reported in this thesis was carried out in the Department of Economics of the Research School for Pacific and Asian Studies of the Australian National University. Unless otherwise indicated, this research is my own work.

Prem J. Thapa

Prem Jung Thapa

June 1999

ACKNOWLEDGMENTS

I gratefully acknowledge the support and encouragement received from the many individuals and institutions, without which it would have not been possible to complete this thesis. I owe a very special debt of gratitude to Professor Peter Warr, my supervisor, for the academic and personal support I received over the period of conceptualizing and completing this research. I received a lot of time and attention from Peter during the course of this study. In addition to his direct contribution in helping me to sort out the analysis of this thesis and through his critical reading of the early drafts, I also learned a great deal from him about the method of economic research and the value of framing the research question precisely.

I have received invaluable advice and help from the other two members of my supervision committee, Dr. George Fane and Dr. K Kalirajan. I wish to acknowledge their contribution in nurturing my research and in improving the analysis and presentation through their comments and suggestions on drafts of my thesis chapters. Others who have kindly read my drafts at different stages and provided helpful comments are Chris Manning, Prema-chandra Athukorala and Xin Meng. I have also benefited from consultations and advice received on specific aspects of my research from Prof. Adrian Pagan, Dennis Hare and Dale Squires. I thank all of them, but remaining errors and omissions are my responsibility.

I also gratefully acknowledge the support and encouragement I have received from Prof. Warwick McKibbin, Head of Department, Prof. Hall Hill and Dr. Ric Shand of the Australia-South Asia Research Center at the ANU.

I am also grateful to Carol Kavanagh for her substantial editorial assistance. For statistical assistance I am grateful to Yvonne Pittelkow, and for IT support, to David Dinham and Roshan Bandara.

I also wish to record the considerable personal assistance and encouragement I received from the management and colleagues at the Institute for Integrated Development Studies (IIDS), Kathmandu, from where I have been on leave during the period of my ANU enrollment. I wish to record a special word of gratitude to Mr. Kul Shekhar Sharma, Dr. Meena Acharya and Dr. Mohan Man Sainju for facilitating my ANU enrollment and the completion of this thesis.

I am very grateful to Nepal Rastra Bank for giving me access to and help in managing the household level data files from the Multipurpose Household Budget Survey of Nepal, that I have used in the empirical analyses of this thesis. A very special thanks goes to Mr. Bhaskar Risal, who I have bothered for help with the survey data and codebooks even years after his official position as Project Director of the survey had elapsed. I would not have been able to perform the estimation work of this thesis without his unrelenting support.

Some sections of the Rastra Bank survey data were also processed at IIDS from the original survey questionnaires. I am grateful again to IIDS for organizing and financially supporting this additional data processing. Many individuals contributed to this work. I wish to express my gratitude in particular to Dr. Krishna Dev Bhatta, Asosthama Pokhrel, Shanker Aryal, Gajendra Shrestha and Ram Kumar Biswokarma and his team of data entry operators.

I have received financial support from many sources during the period of my Ph.D. studies at the Australian National University. I gratefully acknowledge the support provided by the World Bank Graduate Scholarship Program, the Ph.D. Scholarship of the Australian National University, and by the Ford Foundation, New Delhi for the payment of my overseas tuition fees for the later part of my ANU enrollment. The Ford Foundation assistance was arranged through a grant made to IIDS, for which I am also grateful to IIDS. I also received additional financial assistance from the Department of Economics, RSPAS/ANU for which I thank Warwick McKibbin, Peter Warr and Hal Hill for their contributions.

The findings of this thesis have been presented in the regular student seminar series of the Economics Department/RSPAS and the National Center for Development Studies. The valuable comments and suggestions I received in these seminars from Neil Vousden, other departmental staff and fellow Ph.D. students are all gratefully acknowledged.

I also wish to record my appreciation of the moral support and friendship offered by my fellow Ph.D. students, Agus, Harley, Kelly, Salim, Alan, Xialou, Renuka, Iman, KK, Lal, Kanhaiya, Alexandra, Hsiao-Chuan and Xiaoqin and. A special word of thanks to Krishna Hamal and H. M Pant from the small but very supportive Nepali community here in Canberra.

My deep gratitude goes to my wife, Bijaya, for love, support and the admirable patience that she has shown during the long period of this study. I have been on borrowed time that rightfully belonged to her and the children, Siddhanta and Proshanta. I also acknowledge the personal support received from the extended family in Nepal, especially from my parents and from San Dijju and Bhinaju, and thank them for the special assistance and long distance care received during our Canberra sojourn. Finally, I also gratefully acknowledge the cheerful encouragement and appreciation of my research work that I have received from my uncle, C. S. Thapa, from which I drew much of the inspiration needed to complete this thesis.

Prem J. Thapa

ABSTRACT

Is hired labour a perfect substitute for family labour as inputs in farm production? Are these two types of labour equally productive? A negative answer to either question has important implications for the analytical modeling and empirical estimation of the labour demand and supply decisions of farm households, and for deriving welfare measures based on the implicit valuation (shadow price) of family labour. Significant policy implications also arise with imperfect substitutability and/or productivity differences because aggregate outcomes will be sensitive to changes in the distribution of individual endowments of land and labour.

This study uses a production function based approach to test for heterogeneity between family and hired labour in crop production in the *tarai* (southern lowland) region of Nepal. It develops an appropriate analytical framework and carries out the empirical estimation of an aggregate farm household model that allows for the heterogeneity between family and hired labour. A sequential estimation strategy is adopted. The labour heterogeneity detected in the first step of the production function estimation is incorporated, at the second step, in the labour supply estimation in a theoretically consistent manner. The methodological novelty is to relate the shadow wage rate for family labour to the observed market wage rate for hired labour. This is done on the basis of the differential productivity of family and hired labour detected in the production function estimates.

The production function estimation results, based on a translog equation, indicated that although family and hired labour are perfect substitutes, they are not equally productive. A linear labour aggregator function, with a constant efficiency difference between family and hired labour, is the preferred specification to describe the nature of labour heterogeneity. When measured in *effective* units, one unit of hired labour substitutes for 0.75 units of family labour. This efficiency difference is statistically significant and robust to alternative specifications and assumptions on parametric restrictions.

The lower efficiency of hired labour, which is independent of other input levels, implies that farm households should not generally engage in simultaneously hiring in and hiring out of labour. This prediction is indeed borne out by the sample data from Nepal.

The linear form of the efficiency difference between family and hired labour maintains the recursive structure of the conventional farm household model structure. However, the effective wage rate for family labour differs according to the market exposure of the household in the hired labour market. In the labour supply estimation results, model specifications that account for the wage gap between households that are net sellers and buyers of labour are preferred to specifications that assume a common wage rate for all households. This result provides an independent corroboration of the efficiency difference between family and hired labour derived from the production function.

The factor demand elasticities and labour supply elasticities, derived from the farm household model with labour heterogeneity, are all within reasonable bounds and have expected signs. Accounting for the efficiency related wage gap is important in the labour supply estimation because it leads to differences in the elasticity of male and female labour supply with respect to wage, although these differences are not dramatic.

The efficiency difference between family and hired labour provides an alternative explanation for the relatively lower per hectare input of labour on the bigger farms, without relying on labour or other factor market failures. It would be very useful to direct further research to discriminate between alternative explanations for the underlying source of the labour heterogeneity so that firm implications for land reform and rural labour market policies can emerge.

Table of Contents

LIST OF TABLES

Chapter V

Appendix Table 5A.1

Chapter VI

Appendix Table

Chapter VII

LIST OF TABLES (contd.)

LIST OF FIGURES

LIST OF ABBREVIATIONS

AES	Allen partial elasticity of substitution
AIDS	Almost Ideal Demand system
CBS	Central Bureau of Statistics, Kathmandu, Nepal
CES	constant elasticity of substitution production function
CRTS	constant returns to scale
CRINT	cropping intensity
CSS	complete strong separability of inputs
DES	direct elasticity of substitution
DF	degrees of freedom
GL	generalized linear production function (Diewert)
HEC	Hicksian Elasticity of complementarity
HMG/N	His Majesty's Government of Nepal
ICRISAT	International Crop Research Institute for the Semi-Arid Tropics
IMR	inverse Mill's ratio
MPHBS	Multipurpose Household Budget Survey (conducted by Nepal Rastra Bank)
MRS	Marginal Rate of Substitution
NL	Non-linear least squares
NRB	Nepal Rastra Bank
OLS	ordinary least squares
PSS	partial strong separability of inputs
TL	Translog production function
VES	variable elasticity of substitution production function
WS	weak separability of inputs in production function

GUIDE TO NOTATION

A	land input (generic)
Ac	net farm cultivated area
Ae	effective land input
Ag	gross farm harvested area
B	bullock power input
C	composite consumption good in the household utility function
E	pure non-labour income of farm household, excluding farm profit
E*	non-labour income of farm household, including farm profit
F	family labour work-days allocated to own farm-cultivation
H	hired labour work-days allocated to farm cultivation
$\mathcal{L}$	leisure in the household utility function
Ld	labour demand
Le	effective labour input in farm cultivation
Ls, L	labour supply
M	**in Chapter VI** : material inputs in crop production (one of four main inputs specified in the translog production function)
M	**elsewhere** : family labour days worked in the off-farm labour market for a wage
NLY	household level non-labor income
p	price of farm output
π	farm profit : pure return to farm household from land ownership
p_c	price of consumption good
PNLY	per capita non-labor income
Q	aggregate farm output quantity
θ	ratio of relative productivity of hired to family labour in linear labour aggregator function
θ*	ratio of the marginal product of hired labor to family labour at the optimum labour allocation in farm cultivation
T	total time endowment
w	wage rate received by family labour for off-farm labour supply

CHAPTER I

INTRODUCTION

1.1 Motivation

Family based agricultural/farm households are the principal form of economic organization in most developing countries. Their economic decisions on farm production, labour use and consumption and leisure choices can be quite complex because the farm/household acts both as a producer and consumer unit. In conventional economic analyses production and consumption decisions are treated independently. A firm's decisions on optimal input use are not affected by the utility maximizing consumption decisions of the owners of the firm. But in the traditional agriculture setting, where production inputs are primarily supplied by the farm family, household preferences over the consumption bundle, including leisure, can directly affect production decisions. Conversely, production technology and outcomes can directly affect consumption choices, in addition to the conventional income effects.

When decisions on production and consumption are systematically inter-linked, behavioural responses to output and factor price change, and to other policy interventions, can be quite different from those suggested by conventional analysis. Hence, detailed empirical models of the agricultural household that recognize its dual role as a producer and consumer unit in a theoretically consistent manner have become essential tools for policy analyses.

The methodological approach to analyzing farm-household decision making that explicitly accounts for the inter-relationships between production, consumption and labour supply decisions is referred to as a farm (or agricultural) household model. The empirical application of this approach was pioneered in the late 1970's, building on the theoretical work on the "subjective equilibrium" of the family farm household developed by Nakajima (1969) and Sen (1966), among others.[1] There is now a large literature on the estimation of farm household models for developing countries and their empirical applications for a wide range of policy questions.[2]

A key feature of most empirical applications of farm household models is that the many kinds of human labour inputs observed in traditional agriculture are simply added together into an aggregate homogeneous labour input category. The aggregate labour input is then assigned a "price" equivalent to the prevailing market wage rate at which labour can be sold on the local hired labour market. In some few cases the relative productivity differences of labour categories (e.g., male and female, or adult and child) may be indirectly incorporated by creating an aggregate labour input with fixed conversion factors. For instance, when the labour input data is dis-aggregated by gender, a common technique is to create an aggregate labour input by specifying that one unit of female labour is equivalent to, say, 0.8 units of male labour. Such conversion factors can be *ad hoc* or can be based on the ratio of the observed market wage rates for male and female labour.

Another important dimension to the dis-aggregation of total farm labour input is the distinction between labour supplied by family members of the farm households and the labour input provided by labourers hired at a given wage rate. The distinction between *family* and *hired* labour in farm household models is relevant not just from an empirical perspective of capturing any productivity differences between these two types of labour inputs. This distinction is also important for methodological reasons also because it determines the analytical structure of the farm household model and the manner in which the household equilibrium can be determined.

---

[1] The initial theoretical reference to farm household models was a 1923 book by the Russian economist, A. V. Chayanov, of which an edited English version was published as Chayanov (1966). There were also several important non-English publications of Japanese economists - Tanaka (1951), Nakajima (1949). Early examples of the empirical estimation of farm household models are Lau, Lin and Yotopoulos (1978) for Taiwan, Kuroda and Yotopoulos (1978, 1980) for Japan and Adulavidhaya, et.al. (1979) for Thailand and Barnum and Squire (1979) for Malaysia.

[2] Two important collections of theoretical and applied work on farm household models can be found in the edited volumes Singh, Squire and Strauss (1986) and Caillavet, Guyomard and Lifran (1994).

In the classification scheme adopted in Nakajima's classic treatise (1986) on farm household models, the ratio of family and hired labour use is one of the two key dimensions according to which farm households can be classified.[3]

At one extreme are settings where no local hired labour market exists and all farm cultivation is done by family labour. In this situation, where the household is in a state of "autarchy" (i.e., neither buys nor sells any labour), the household labour supply and labor demand equilibrium is characterized by a "subjective equilibrium" (Nakajima, 1986). Such an equilibrium implicitly defines an unobserved "virtual" or "shadow" price of labour, which, if it existed in reality, would be the price (the wage rate) that would induce the household to equate the demand and supply for its own family labour, taking that implied market wage rate as exogenously given (Strauss 1986:77). The key analytical feature of the autarchic farm household is that its preferences over leisure and consumption of goods, together with the farm production function, jointly determine the optimal demand for labour in farm production. Optimal production and consumption plans are formulated simultaneously.

At the other extreme are farms that rely completely on non-family labour employed at a specified market wage rate. These are "farm firms". The decisions on labour allocation for farm cultivation (and other optimal input choices) are exactly analogous to the textbook version of the profit maximizing firm. The consumption choices of the farm-firm household can also be modeled as in the textbook version of a consumer household. It chooses an optimum consumption bundle, given its budget constraint (which in this case includes income from farm profits) and exogenously given market prices, including the market wage for hired labour as the appropriate price of leisure. The only special condition to take account of is the revealed preference for zero amounts of family labour supply. The non-labour participation of the household members indicates that the subjective valuation placed

[3] The other dimension is the proportion of farm output consumed by the farm household. Various categories with a distinct analytical structure of the farm household model result from their different exposure to the hired labour and farm output markets (Nakajima 1986, Chapter 1).

on family leisure must exceed the market wage rate at which family labour could be sold. Such non-participation kinks are readily handled in labour supply models (Killingsworth 1983). The production decisions of the farm-firm household are separable from the consumption choices.

The vast majority of farm households in both developing and developed countries fall into the intermediate category where both family and hired labour are used in farm cultivation. In addition, with active rural labour markets, farm households have the option of allocating family labour to own farm cultivation or to wage labour employment in the local off-farm labour market.

The conventional approach to empirical estimation of a farm household model which has both family and hired labour inputs has been to treat family and hired labour as homogeneous inputs which can be substituted for each other on a one for one basis. Total labour input on the farm is simply taken to be the sum of family and hired labour days. It is also widely assumed that the local (or village) wage rate at which labour can be hired in is also the appropriate wage rate applicable to family labour. Such an imputation is often made even if family labour is fully devoted to own farm work and does not supply any labour on the hired labour market (Rosenzweig 1980).

These twin assumptions about the equivalence of a unit of family and hired labour (homogeneity) and an inferred, exogenously given market wage rate for family labour have the effect of making family labour equivalent to a tradable input with a price that does not depend on the farm-family's own decisions. If all other farm inputs and outputs are also tradable commodities with exogenously given market prices that the farm household takes as given, then the analytical structure of the farm household model is considerably simplified. The production and consumption decisions need not be modeled *jointly* (which makes empirical estimation relatively difficult) but can be solved *recursively*, so that the production and consumption decisions can be determined separately.

In the recursive framework, the farm-household utility maximization problem is solved in two stages. In the first stage, the household acts as a profit maximizing producer which chooses optimal levels of farm inputs (including total labour which is a homogeneous composite of family and hired labour), given the exogenously given prices for all inputs and outputs that it faces. The solution to this first step determines farm profits and the full income of the farm household. Given this income level, in the second stage the farm household acts as a pure consumer, choosing the optimal consumption bundle, including leisure, subject to the full income budget constraint conditioned on the optimal production side choices. A simple interpretation of the recursive property of the farm household model is that the profit maximizing production input and output choices are independent of the household's utility function (Strauss 1986).

An intuitive way to understand the recursive property is that it necessarily holds as long as there are markets for all commodities and inputs and the household is a price taker in all these markets. Then the amount of, say, the food crop to be produced by the household can be determined independently of the preferred level of household food consumption, because the household can always buy and sell any amount of food at the going market price. Similarly, household preferences about leisure relative to the other consumption goods together with the market wage determine labour supply independently of the labour requirement on the production side. Again, any difference between household labour supply and on-farm labour demand can be costlessly equilibrated through hiring in outside labour or hiring out family labour at the going wage rate.

Although the treatment of family and hired labour as homogeneous inputs greatly facilitates empirical estimation of farm household models, there are important reasons for carefully assessing whether family and hired labour are indeed homogeneous inputs in a given setting.

The main reason why family and hired labour may differ in terms of efficiency units is due to the "principal-agent" type of relationship inherent in most hired labour contracts (Binswagner and Rosenzweig 1986). Because of the difficulty inherent in monitoring the effort applied by hired labour with a limited stock of family labour available to the farm household "shirking" on the effort applied by hired labour may be commonplace (Feder 1985). On the other hand, family labour has an incentive to work intensively because on the owner operated family farm it is the residual claimant to output. Many of the institutional structures of traditional agriculture - for instance, multi-generation joint families, sharecropping contracts, piece-rate wages, labour exchanges - seem designed to get around the problem of monitoring the work effort of hired labour.[4] Hence there is implicit recognition that the efficiency differences between family and hired labour related to the effort applied in farm work can be large.

Apart from the implications for an analytical understanding and modeling of the labour demand and supply decisions of farm households, any observed efficiency differences or heterogeneity between family and hired labour can also have significant policy implications. With efficiency differences between family and hired labour related to work incentives, aggregate outcomes - e.g. total production, total labour demand (labour absorption), and equilibrium rural wage rates - are sensitive to changes in the distribution of individual endowments of land and labour.

If, for instance, family labour is more productive because it applies more *effort* per unit time than hired labour, then a re-distributive land reform program which transferred land from big farms relying primarily on hired labour to small family labour-operated farms would increase the average labour intensity of cultivation, and hence total absorption of labour in the agricultural sector. On the other hand, the *increase* in aggregate labour demand could be associated with a *reduction* in the demand and supply for hired labour such that the equilibrium market wage rate for hired labour could be reduced. This would adversely affect landless households which were not beneficiaries of the land transfers. The direction and exact

[4] In broader terms these arrangements in traditional agriculture can be seen as an institutional response to the high likelihood of incomplete or missing markets, particularly with respect to labour (de Janvry, Fafchamps and Sadoulet 1991).

magnitude of such general equilibrium effects depend on the precise values of the various elasticities of labour demand and supply for the different classes of farm households (Rosenzweig 1978). Hence, from a policy perspective, it is useful to devise a suitable framework and methodology to estimate the parameters of a farm household model that allows for the heterogeneity between family and hired labour.

The empirical literature on testing whether family and hired labour are heterogeneous inputs, and, if so, what determines their relative efficiencies, is rather scant. Moreover, existing studies mainly look at the heterogeneity and efficiency issue solely from the production function perspective, without relating the heterogeneity to the labour-leisure utility maximization problem faced by the household in the presence of labour heterogeneity. Examples of the production function approach are Bardhan (1973), Deolalikar and Vijbergen (1987) and Frisvold (1994). These previous studies on the efficiency differences between family and hired labour have been limited to a production focus only. They do not incorporate an integrated a farm household model structure where the labour supply implications of the observed heterogeneity have also been derived in a consistent manner and verified empirically, as is done in the approach taken in this thesis.

1.2 The Research Question and Methodological Approach

The specific research question addressed is whether, in the setting of Nepalese agriculture, family and hired labour are equivalent (or homogeneous) inputs in farm production. If not, what is the nature and extent of heterogeneity between family and hired labour? And how can one amend the conventional recursive farm household model structure to estimate a theoretically consistent labour supply equation for family labour, taking account of the specific form of the heterogeneity between family and hired labour detected empirically.

There are two distinct though related questions regarding family and hired labour being homogeneous inputs:

(i) Are family and hired labour perfect substitutes for each other in the production function? Is the elasticity of substitution between family and hired labour infinitely large?[5]

(ii) If so, are the marginal products per unit of labour time of these two types of labour equal to each other, everything else held constant? [6]

If the answer to either question above is no, then family and hired labour become heterogeneous inputs. The total effective labour input on a farm utilizing both family and hired labour work-days will be something different than the simple sum of the labour days put in by family and hired workers. More importantly, with labour heterogeneity the wage rate for family labour that applies to the determination of the farm household's labour-leisure equilibrium will differ from the market wage rate paid for hired labour. With this "shadow" wage rate for family labour being unobservable, as well as possibly endogenous to the household's labour allocation decisions, the recursive or separable property of the farm household model is lost and estimation becomes cumbersome.

In the presence of labour heterogeneity, the analytical challenge is to seek ways in which separability could be restored in a theoretically consistent way by relating the "shadow" wage rate of family labour to the observed market wage rate for hired labour, taking into account the precise nature of the efficiency differences between family and hired labour. If the appropriate shadow wage rates for family labour can be recovered from the observed market wage rates and from the fixed parameters of

[5] In a production function with more than two inputs there are many alternative measures of the degree of substitution between two specific inputs (Chambers 1986, ch.1). A commonly used measure is the partial Allen Elasticity of Substitution (AES) which gives the effect on the quantity demanded of one factor due to a change in the price of another factor, holding output and other factors constant (Squires 1994:186). A positive value of the AES implies the two inputs are substitutes, or more formally, price or *p*-substitutes (Seidman 1989). The higher the value of the AES the greater the degree of substitutability. For perfect substitutes, the AES is infinitely large. This implies a linear aggregate labour composite of the form $Le = aF + bH$, where Le is aggregate labour, F is family and H is hired labour, and a, b are constants.

[6] Since marginal products of inputs vary with the levels of other inputs, the comparison of the marginal products of family and hired labour has to be made at some average level of all inputs in the sample of farm households. But if the production function is separable in the labour inputs, in which case the ratio of the marginal products of family and hired labour depend only on the levels of the labour inputs (Chambers 1986:43), the comparison of marginal products can be made only in reference to some average level of the labour inputs. See Chapter III for details.

the production function which describe the extent of the efficiency differences, then separability of the farm household model is restored. The labour supply component can be estimated with the shadow wage rate generated in this manner. Whether such an approach is feasible will depend on the type of the labour heterogeneity indicated in the data and on how the marginal products of hired and family labour behave.

An important limitation of this thesis needs to be noted at the outset. It does not delve into the underlying sources of the efficiency differences between family and hired labour implied by the production function estimation results presented in Chapter VI. This research is not intended to discriminate between alternative explanations for such differences may arise - e.g., whether they are due to the principal-agent incentive problem for hired labour, or the farm-specific experience of elderly family members (Rosenzweig and Wolpin 1986), or some undetected labour market imperfections. That would be an altogether different thesis, and one which cannot be adequately supported by the type of survey data used in this study.

Estimation methodology

The empirical estimation strategy adopted is based on a two-step procedure suggested by H. Jacoby (1993) for estimating non-separable farm household models.[7] In the first step Jacoby estimates a farm production function with family labour as a distinct labour input, and in the second step a structural labour supply equation (which represent the household's utility maximization equilibrium over leisure and goods consumption) is estimated. The labour supply equation is based on an unobservable shadow wage rate for family labour. Relying on the optimal first order conditions, Jacoby uses the estimates of the marginal product of family labour derived from the production function parameters to identify the unobservable shadow wage rates.

[7] Jacoby's approach is a general methodology for estimating a structural labour supply equation for workers who are self-employed. It is analogous to the treatment of labour supply in the presence of progressive income taxes that was pioneered by R. E. Hall (1973) and which involves "linearizing" an underlying non-linear budget constraint.

Jacoby's prime interest is not to test for efficiency differences between family and hired labour. The same sequential approach, however, can be used to estimate a farm household model structure in which is embedded a test for the heterogeneity of family and hired labour. In the second step the labour supply functions are estimated in a manner which is theoretically consistent with the type of labour heterogeneity indicated in the production function estimation.

The heterogeneity of family and hired labour is tested through several alternative specifications of the production function that allow for imperfect substitution between family and hired labour and for differences in their marginal products. The rejection of the parametric restrictions that lead to a model with homogeneous labour inputs is statistical evidence for labour heterogeneity.

In the second step, some of the variables used in the labour supply regressions will be generated from parameters estimated in the production function.[8] The main variable of interest derived in this way is the appropriate "shadow" wage rate that reflects the true opportunity cost of family labour at the equilibrium labour supply position. A key question is how this shadow wage rate is related to the off-farm market wage for family labour and the wage for hiring in labour on the farm.

The first order conditions for equilibrium of farm household in a model with heterogeneous labour inputs can be used to relate the shadow wage rate to the observed market wage rates and the parameters which describe the extent of the labour heterogeneity. The labour supply equations can then be estimated with the appropriate shadow wage rates derived in this manner. The resulting labour supply regression parameters then describe the true labour/leisure choice of the farm household that is consistent with utility maximization in a setting where the specific type of labour heterogeneity modeled in the production function is observed.

---

[8] The sequential estimation of models that contain variables that are unobservable but can be estimated from an auxiliary statistical model is a standard procedure. Early examples from many areas of applied econometrics are cited in Murphy and Topel (1985). Pagan (1984) refers to these as models with "generated regressors". The sequential estimation strategy yields consistent estimates of second stage parameters under fairly general conditions, requiring only a simple adjustment of the standard errors for the second step equation (Pagan 1986 and Murphy and Topel 1985).

1.3 The Setting

The data used in the empirical section of this thesis comes from a sample of farm households from the southern plain (or *tarai*) region of Nepal. The household level survey data used for the estimation are drawn as a sub-sample of a nationally representative Household Budget survey conducted by Nepal Rastra Bank (the Nepalese central bank) in 1985/86. This survey had a farm management module in which data on output and input use was collected on a crop basis over a multi-round survey period. The data separately identifies inputs of family and hired labour, which are further broken down into male and female categories. This survey also included a module on family labour use by household members aged ten or above, with which labour supply estimation can be done at the individual level.

The sample size of the data used for the estimation of the farm household model is 1007 households drawn from five of the twenty districts of the *tarai* region of Nepal. (See Figure 1.1). Data for a larger sample of households from the northern hill and mountain regions of Nepal is also available in the 1985/86 Nepal Rastra Bank survey. But these households have been excluded from this study because of the very limited use of hired labour in the northern hill and mountain regions of Nepal. Farm sizes in these northern regions are very small, production is mainly subsistence oriented, and there is only a limited form of economic differentiation among farm households in this region.

The agrarian structure in the southern plain region of Nepal more closely follows the classical structure in which most villages will have a few large landlords dependent primarily on hired labour, a middle group of owner-cultivators and a large group of landless households who supply the hired labour on the larger farms and may also typically rent in small plots of land. Hence, in the *tarai* region it becomes feasible to distinguish between households that are net buyers and sellers of labour on the local village labour markets.

The rural labour markets in the *tarai* region are quite active. In peak periods there is also considerable in-migration of agricultural labourers from India. Rural households in the *tarai* region of Nepal are less likely to be constrained quantitatively on either the labour demand or labour supply side than farm households in the northern hill/ mountain regions. Hence the modeling of efficiency differences between family and hired labour and the implied correctly specified labour supply behaviour is best addressed in the context of the *tarai* region only.[9]

The productivity of agricultural labour is a central theme in any policy discussion of rural development and poverty alleviation in Nepal. Agriculture still accounts for about 60% of GDP. About 90% of the Nepalese live in rural areas and more than 80% of them depend on agriculture for their livelihood. Agricultural production technology is still very backward in Nepal. Human labour is the main farm input, typically accounting for more than 50% of the total cost of farm cultivation.[10]

The parametric estimation of a farm household model structure for Nepal with a specific focus on labour productivity is useful work in itself, apart from the methodological concerns of separability under labour heterogeneity. There has been little quantitative work of this nature on Nepalese agriculture utilizing large sample surveys. Obtaining more precise estimates of these parameters - for instance, labour supply and factor demand elasticities - would be valuable inputs for the analyses of many agricultural policy issues in Nepal, including applications of general equilibrium simulation models.

1.4 Thesis Outline

Chapter II provides a brief discussion of the background issues related to the question of labour heterogeneity. It also reviews the literature on the methodology as well as the findings on the tests of the heterogeneity of family and hired labour

---

[9] A comparative analyses of farm household behaviour and model parameters for the *tarai* and northern hill regions of Nepal would be of interest, but this has been left for subsequent research.
[10] See *Costs of Production for Major Crops in Nepal 1985/86* (Ministry of Agriculture, Kathmandu).

carried out in previous studies; and on related aspects of the larger literature on labour supply behaviour of farm households that are relevant for this study.

Chapter III develops the analytical framework of a farm household model with heterogeneous labour inputs and the labour supply estimation implications of specific forms of heterogeneity. This chapter also shows how the correct specification of the labour supply equations differs among households in the presence of labour heterogeneity.

Chapter IV develops and discusses the specific econometric estimation methodology adopted in this thesis. The approach essentially follows the sequential two-step procedure proposed by Jacoby (1993) with a specific adaptation that relates in a theoretically consistent manner the variables in the labour supply estimation to the type of labour heterogeneity that was detected in the production function estimation.

Chapter V briefly describes the Nepalese agriculture setting from which the survey data used in this thesis is drawn. It also provides an overview of the survey data itself and the derivations and definitions of variables used in the estimation chapters.

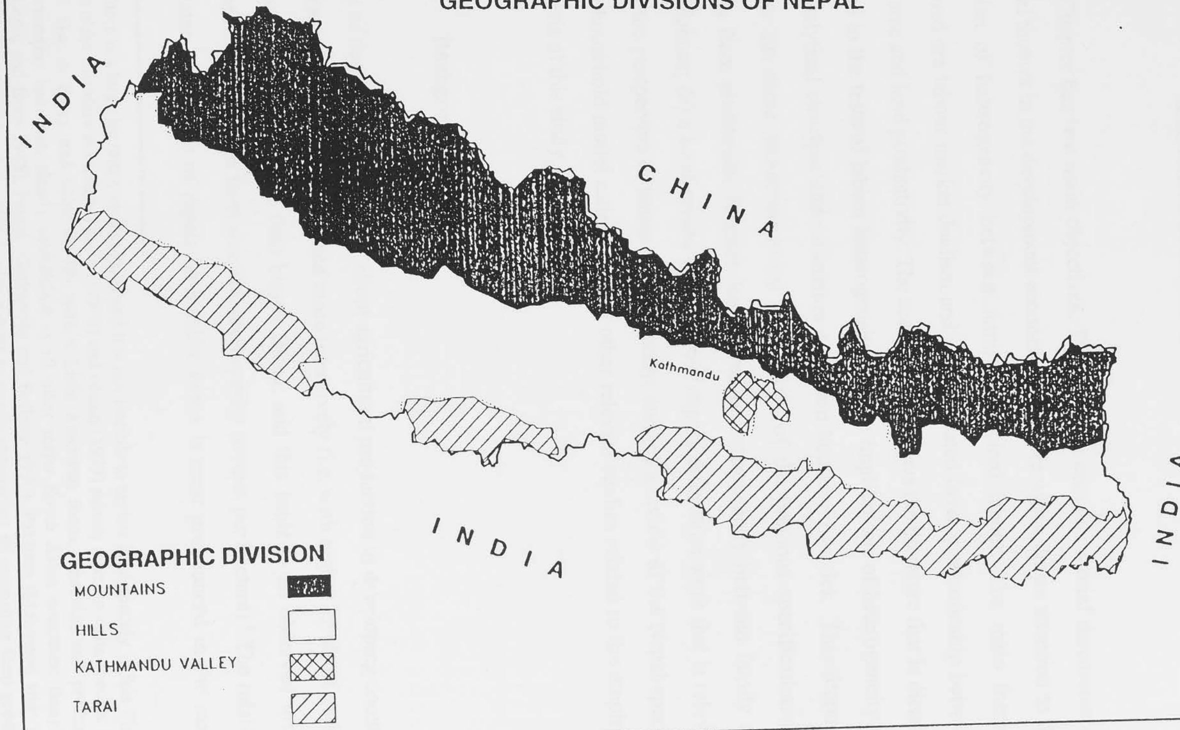
Chapter VI contains the estimation results on the tests for labour heterogeneity and alternative aggregator functions to create an effective labour composite combining family and hired labour. It also discusses the test for separability of the labour input from other inputs, as is required in order to create consistent labour aggregates.

Chapter VII presents the labour supply estimation results consistent with the nature of labour heterogeneity detected in Chapter VI. It also compares the model fit and parameter estimates of alternative specifications of the labour supply models with and without the heterogeneity adjustments.

Chapter VIII contains the summary and conclusions. It also briefly draws out some implications of the research and provides suggestions for future work.

FIGURE 1.1

GEOGRAPHIC DIVISIONS OF NEPAL



CHAPTER II

BACKGROUND AND LITERATURE REVIEW

This Chapter has two main objectives. The first part contains a brief discussion of several themes in the development economics literature which draws attention to the question of heterogeneity between family and hired labour. The main themes explored are labour market dualism and the celebrated inverse relationship between farm size and land productivity. The second part reviews the literature that is directly related to the tests of labour heterogeneity and the implications of heterogeneity for the analytical structure and estimation of a farm household model. This literature review has three sub-sections: (a) a summary of the previous specification and results from production function based tests of heterogeneity between family and hired labour; (b) a brief review of the labour supply estimation work that is relevant from the perspective of labour heterogeneity; and (c) a review of the Nepal-specific farm household model estimation and other relevant studies related to the empirical analysis of this study.

2.1 Background Issues

One of the key "stylized facts" about agricultural production in developing countries is that small farms are cultivated more intensively (i.e. with higher levels of variable inputs used per hectare) than bigger farms; and this leads to an observed inverse relationship between farm size and productivity (output per hectare).[1] The relatively greater application of inputs on smaller farms is most pronounced in the case of

[1] There is a large literature on this topic and it is not feasible to review it adequately in this Chapter. The older literature is surveyed in Sen (1975) and Ghosh (1979) mainly from an India-specific focus, and also in Berry and Cline (1979) with a Latin American focus. The size and productivity relationship has been clearly established in all other major South Asian countries: Bangladesh (Abedin and Bose 1988), Nepal (Grabowski and Belbase 1990), Pakistan (Mahmood and Haque 1981). The inverse relationship is also robust to newer approaches to estimating farm production functions, including the random coefficient method (Hoque 1988). For a dissenting view on the genuine nature of the inverse relationship see Rudra (1992).

labour inputs. The labour input per hectare on small farms is consistently higher than on bigger farms over a large range of farm sizes; and this result holds whether or not average yields on small farms are higher (Berry and Cline 1979).

Based on the above stylized facts of traditional agriculture, many authors went on to claim that a re-distributive land reform policy would be desirable not only from an equity point of view, but also from an efficiency consideration because breaking up hired-labour-based large farms into smaller family based farms would also increase agricultural output (Berry and Cline 1979). Such a line of *ceteris paribus* reasoning can be quite misleading. Everything else is unlikely to be the same after the land reforms. The early literature also did not address precisely what were the sources of the labour market dualism leading to higher intensity of cultivation on the smaller farms, and whether they would still be maintained after the land transfers.

Traditionally three general explanations have been offered for the observed inverse relationship, with regard to both labour input and yields:

a. decreasing returns to scale;
b. omitted variable bias due to unobserved higher quality of land on small farms; and
c. factor market imperfections, especially in rural labour markets, that induce small farmers to apply too much family labour on their own-farm cultivation, relative to the market wages they face.

Of these three explanations, the empirical evidence for decreasing returns to scale in traditional agriculture is quite weak (Bardhan 1973, Moll 1990). Also, it is not a very plausible explanation in settings of extremely small average farm sizes. For instance, Benjamin (1995) documents a strong inverse relationship in Java within a sample of farms where the mean farm size is 0.71 hectares with a standard deviation of 0.01. There would be few *a priori* reasons to expect decreasing returns to scale in the technology available to farmers within such a narrow distribution of farm size.

More recent work on the inverse size and productivity relationship has focussed on land quality heterogeneity (Bhalla and Roy 1988). This new evidence, especially from India, indicates that smaller farms tend to be more productive because of better irrigation facilities and other innate land quality differentials that in turn induce the higher labour input. In such a setting, breaking up the big farms into many smaller ones would not have any output increasing effect.

The evidence is mixed about the relative importance of these two main explanations for the inverse relationship - unobserved land quality differences or factor market imperfections. Where the data allow for controlling the effects of possible correlation between farm size and unobserved land quality, the inverse relationship with respect to output sometimes disappears, as in Bhalla and Roy (1988) with Indian data, or still remains strong, as in Carter (1984), also with Indian data.

The traditional explanations focussing on labour market imperfections were based on a hypothesized difference in the "mode of cultivation" between small and big farms, leading to a form of labour market dualism within the agricultural sector. The causes and consequences of labour market dualism in developing countries has been an important strand in the development economics literature for a long time. While inter-sectoral wage gaps (i.e. a gap between the rural agricultural labour wage rate and an urban sector labour wage rate in real terms) have been the main focus of this literature, following Harris and Todaro (1970), the presence of dualism within the rural/agricultural sector is also widely recognized (Mazumdar 1975, 1989; Iqbal 1981).

In a farm household model framework, the essence of labour market dualism lies in the simple presumption that the real cost of unpaid family labour - measured as the rate of utility substitution between leisure and consumption - is not equated to a common market wage rate for labour (Sen 1975). Consequently, the real cost of (identical) labour can differ among rural households even in a localized village labour market.

This line of explanation, while plausible in the institutional setting of traditional family based agriculture, is not wholly consistent with the empirical facts that small farms also rely to a considerable extent on hired labour; and that a large part of the hired labour on big farms is supplied by self-cultivating small farmers. In addition, there is a growing general acceptance and empirical support for relatively well functioning rural labour markets (Rosenzweig 1986, Benjamin 1992) which goes against the main grain of Sen's explanation. One should note that the existence of some form of labour market imperfections also implies that there are related imperfections in the land rental markets that prevent the wage gap being erased through changes in the scale of operation of the small and big farms (Binswagner and Rosenzweig 1986).

An extreme form of the inverse relationship between farm size and productivity occurs when the output per hectare is strongly declining in the size of a plot within the same farm-household. Udry, *et al.* (1995) estimate a farm production function using the very detailed plot level data collected by ICRISAT from Burkina Faso. With a panel data estimation framework they show that, controlling for household and annual fixed effects, output is strongly declining with the size of the plot within a farm. The usual factor market imperfection story cannot explain such a relationship since it occurs among plots cultivated within the same household, and not across households that may face different opportunity costs of labour. Among a variety of possible explanations, Udry, *et al.* note that this may be an indication of labour monitoring problems on bigger plot within a household.

An alternative explanation for the lower labour input per unit of land on bigger farms, which does not rely on labour market failure, is to explicitly account for the differences in the efficiency of hired and family labour time due to difficulties in monitoring the effort applied by hired labour. Feder (1985) has formalized a model with supervisory costs of using hired labour (and with credit constraints) which gives rise to an inverse relationship between farm size and labour use and output, even with well functioning land and labour markets. Incentive related differences in the effort applied by farm labour in other forms of agricultural organizations, such as collective farms, has also been noted (Sicular 1986).

However, there has been very limited work on empirical testing for heterogeneity between family and hired labour. Still less work has been done to show whether any observed heterogeneity can be related to the principal-agent type of incentive problem (Douglas 1989), or to the related costs of monitoring the effort applied by hired labour (Binswagner and Rosenzweig 1986). If labour heterogeneity is related to work incentives, with family labour being more productive, this would offer an alternative explanation for the inverse relationship between labour use and farm size without explicitly relying on factor market imperfections.

Even when there is clear evidence for labour heterogeneity, with family labour based small farms being more productive, the efficiency based argument for land reform is not clearly established without knowing in detail the sources of the labour heterogeneity. It could be that family labour is more productive solely because the family members in households that own or operate some land are better fed than hired workers who usually come from landless households. There is some evidence that nutritional status does affect the effort expended in farm work (Deolalikar 1988). Hence, for example, if nutritional deficiencies are at the source of labour heterogeneity, the incentive related analytical categories of family and hired labour are not relevant. The nutritional deficiencies of the landless workers could be remedied without turning them into owner-cultivators (although doing so through a re-distributive land reform program would also improve their nutritional status and hence productivity). The policy problem becomes one of choosing a first-best option.

2.2 Literature Review

This section provides a more detailed review of the specific methodology and results obtained in previous studies in relation to direct or indirect tests for heterogeneity between family and hired labour. A second sub-section reviews the labour supply estimation literature from the perspective of the implications of labour heterogeneity. A third part of this section reviews the Nepal-specific literature with respect to empirical estimation of farm household production and labour supply decisions.

(a) Labour heterogeneity

The main approach to testing for heterogeneity between family and hired labour has been through production function estimations that rely on alternative specifications of the labour input. Tests for labour heterogeneity in a production function framework ask two different but related questions. First, are family and hired labour perfect or imperfect substitutes in the farm production process as measured by the elasticity of substitution between them? Second, regardless of the degree of substitution between them, are the marginal effects on output from a unit increase in family labour the same as the effect from an unit increase in hired labour? The latter is a test for the equality of the marginal products of family and hired labour from a purely technological perspective.[2]

These tests can be done with family and hired labour specified as two distinct inputs which enter the production function (Jacoby 1993, Squires 1994). Or, the production function can have a nested structure in which, at the first step, a composite labour services variable can be specified as a function of the observed levels of family and hired labour inputs. At a second stage, the composite labour variable enters as a single input in the production function with other non-labour variables (Deolalikar and Vijverberg 1983, 1987 and Frisvold 1994).[3] The former is a more general specification, but it is not always feasible to implement, especially if there are many cases with zero-values for any of the labour categories in the sample data.

[2] This is a measure of the technical efficiency of hired and family labour. If there is a difference in the wage rates to be paid to family and hired labour, then at the optimal labour allocation, the ratio of the marginal product of family and hired labour would, of course, reflect the difference in their relative wage rates.

[3] Testing for labour heterogeneity with a nested production function structure is valid only when the main production function is separable in the labour inputs. Otherwise a consistent labour aggregate is not defined (Chambers, 19868; Berndt and Christensen 1974). Empirical applications of the nested production function structure have not always been based on the required prior test for labour input separability.

Bardhan (1973) is one of the earliest examples of a specific test for labour heterogeneity embedded in a production function estimated with farm level data (from the Indian Farm Management surveys). Bardhan's production function specification was Cobb-Douglas where the ratio of hired labour to total labour entered as a separate input. The value and the sign of the coefficient on the labour ratio variable provide a test for labour heterogeneity. For most of his different samples Bardhan finds that the coefficient for the ratio variable is insignificant, implying homogeneous labour. In the few cases where the labour ratio variable is significant, its sign is positive, implying hired labour is more productive than family labour.

Deolalikar and Vijverberg (1983) have criticized Bardhan's test as being based on an unsatisfactory functional specification, and also for not distinguishing clearly between imperfect substitutability and differential marginal effects. In particular, a Cobb-Douglas specification with the ratio of hired labour to total labour entered as a separate variable implies that hired labour is an essential input. Output is zero if the hired labour ratio is zero; and this can possibly bias the coefficient on the ratio variable to be positive (Deolalikar and Vijverberg 1983:48).

Using aggregate district-level data for India, Deolalikar and Vijverberg (1983) used a more general functional specification to test for labour heterogeneity. They specify a nested production structure in which the overall production function is Cobb-Douglas with a composite labour services input. The composite labour input, in turn, is specified under various alternatives of a CES or Diewert's generalized linear production function. They find strong evidence for limited substitution between family and hired labour. Their estimates of the Allen partial elasticity of substitution (AES) between family and hired labour ranged from 0.6 to 2.4, under different specifications.[4] They also find substantial difference in the marginal product of family and hired labour, with family labour being more productive.

[4] These are the estimates of AES between family and hired labour in the function for the nested labour input. Such an AES is distinct from the AES between family and hired labour in the overall production function (Deolalikar and Vijverberg 1983).

Using conventional F tests on the restrictions on the model parameters and the Davidson and MacKinnon (1981) J test for non-nested models, Deolalikar and Vijverberg conclude that their "best specification" of the agricultural production function is a Cobb-Douglas relationship between family and hired labour. This specification actually transforms the entire production function also into a Cobb-Douglas relationship with family and hired labour as distinct labour inputs.[5] Such a preferred specification is due in part to the nature of the aggregate data which is at the district level, and hence, all sample points have non-negative values of both the family and hired labour input. In the preferred specification the marginal product of family labour is 2.5 times the marginal product of hired labour at the mean of the data. While a finding of a higher productivity of family labour is plausible, such a large discrepancy in the relative marginal effects is implausible. If it is not due to the aggregate nature of the data, which the authors cite as a possible explanation, it is likely to be an indication of the mis-specification of a Cobb-Douglas production function with family and hired labor as distinct inputs.[6]

Aware of the potential bias in the results due to the use of aggregate district-level data, Deolalikar and Vijverberg (1987) made another effort to test for labour heterogeneity using household level data for a sample of Indian and Malaysian farms. The estimation methodology is similar, with a quadratic labour services function nested into a second level Cobb-Douglas production function. Again, they find strong evidence for imperfect substitutability, but the result on the marginal effects are reversed, with an estimated higher productivity of hired labour in both of

[5] A Cobb-Douglas labour service "production function" nested within an overall Cobb-Douglas production function specified in terms of a composite labour variable is equivalent to a single equation Cobb-Douglas specification with family and hired labour as distinct inputs, without any additional parameter restrictions.

[6] The data used by Deolalikar and Vijverberg (1983) for the labour input are not in terms of actual hours of work but in terms of the average number of persons per farm holding in the family labour and hired labour category in a district. If the average days actually worked for each labour category are different then the estimates of the marginal products of labour per work-day will differ. If the number of days worked on average by one family worker is higher than the number of work days of one hired labour, as is to be expected in poor family farms, then the ratio of the actual marginal product per work-day of family and hired labour could be substantially less than the 2.5 value implied in their estimates.

the Indian and Malaysian samples. At the mean of the data, the estimated elasticity of substitution between family and hired labour is 0.68 in the Indian sample, and 1.1 6 in the Malaysian sample. In both cases these estimated values are not significantly different from unity, implying again that a Cobb-Douglas specification for the labour nest would be appropriate.[7]

The ratio of the marginal product of family to hired labour ranges from 0.32 for the Indian sample to 0.78 in the Malaysian sample. Even when family and hired labour are imperfect substitutes, it is difficult to explain why in the Indian sample the marginal product of hired labour is three times higher than the marginal product of family labour (at the mean of the data). To reconcile this with profit maximizing behaviour the shadow wage rate for these two inputs should vary by a factor of three, which is a highly implausible scenario.

Squires and Tabor (1994) use a translog production function specification to test for the degree of substitution between family and hired in a large regionally stratified sample of Indonesian farms. Their estimates of the Hicksian elasticities of substitution[8] between family and hired labour fluctuate in sign, indicating a substitute relationship in some regions and crops, and a complementary relationship in others. Their overall results show a limited extent of substitution between family and hired labour, in particular when measured by the direct elasticity of substitution between them. These values are usually much smaller than one (the implied value from a Cobb-Douglas specification).

Frisvold (1994) approaches the question of labour heterogeneity from the perspective of supervision costs imposed by hired labour. Using a particular interpretation of supervision costs by assuming that monitoring of hired labour is

---

[7] Deolalikar and Vijverberg (1987) do not actually provide estimates for a Cobb-Douglas labour nest because of the problem of zero values for hired labour in a large portion of the sample for India. Ignoring this problem of zero values, the result of a unitary elasticity of substitution means the generalized quadratic equation could be replaced with the Cobb-Douglas form.

[8] The Hicksian elasticity of substitution and other related concepts of input substitutability in a multi-factor setting are defined and discussed in Section 6.3.4 of Chapter VI.

more effective when family members work together with hired labour, Frisvold measures the effect of employer supervision on the productivity of hired labour. If family and hired labour were perfect substitutes in production with no "shirking" behaviour by hired labour, the supervision effect should be insignificant. If supervision effects were present then family and hired labour would be imperfect substitutes with a certain degree of complementary. Increased allocation of family labour, working along side with hired labour and providing supervision, would increase the marginal productivity of hired labour. Using the very detailed plot level production data collected by ICRISAT for their village level studies from southern India, Frisvold reports three major findings:

(i) family and hired labor are imperfect substitutes, and the elasticity of substitution decreases as the ratio of hired to family labour increases;

(ii) the productivity of hired labour increases with the supervision intensity of family labour, but at a decreasing rate; and

(iii) the labour effort expended per unit of time of family labour is higher than that of hired labour, but this differential diminishes with increases in supervision intensity. At high levels of supervision, the productivity of hired labour approaches that of family labour.

Because of the very detailed ICRISAT data set used in Frisvold's study, these results are very significant.

The question of labour heterogeneity has also been addressed in the context of other labour categories, particularly between the labour input of male and female workers. Laufer (1985), using the same ICRISAT data for India, has estimated the elasticity of substitution between male and female labour using a generalized quadratic production function with male and female labour as distinct inputs. She finds evidence for imperfect substitution, but her results, based on separate estimations for different crops, give widely divergent values for the AES. Male and female labour are estimated to be complements (negative AES) in rice and sorghum production but substitutes (positive AES) in legumes. Laufer explains the wide divergence of results by noting that the notional categories of male and female labour may actually be capturing the functional nature of the different tasks performed by male and female workers, given the strict division of labour by gender.

Another important result presented by Laufer is the comparison of the ratio of the marginal product of male and female labour. The marginal product of female labour is always lower than that of male labour. Their ratio ranges from 0.49 to 0.77 and this variation is consistent with the range of the observed ratio of the female to male wage rates in the sample villages.

Udry (1996), using the detailed plot level ICRISAT data for Burkina Faso, also finds substantial gender differences of a more general nature. Using a nested CES production function structure with a multi-factor CES labour nest that specifies four labour categories - adult family male, adult family female, family child and non-household labour - Udry estimates the elasticity of substitution for each pair to be 0.63.[9] Udry also presents dramatic evidence that in a given household, where some plots are under the control of the male husband and other plots under the control of his wives, the productivity on plots controlled by women are significantly less than the productivity on plots controlled by the husband for the same crop. This result is not due to differences in technology or lower efficiency of female labour, which tends to be concentrated in plots controlled by women, but to an overall lower intensity of cultivation of plots controlled by women. Using the same data set, Udry, *et al.* (1995) actually estimate that the marginal product of female labour is higher than male labour over all plots.[10]

---

[9] In a multi-factor CES production function the Allen partial elasticity of substitution between each input pair must be the same (Uzawa 1963).

[10] Udry, *et al.* (1995) acknowledge the higher marginal product of female labour is due partly to crop-composition effects since women tend to specialize in high value vegetable crops grown on their plot with predominantly female labour.

Empirical tests for labour heterogeneity have not always rejected the traditional assumption of homogeneity between family and hired labour. Benjamin (1992), using detailed SUSENAS survey data for Indonesia and concentrating only on rice growing farms, tests for a linear specification of an effective labour composite of the form $Le = L^F + \theta L^H$.[11] He does not reject the null hypothesis of $\theta = 1$, and hence does not reject the joint null hypothesis of perfect substitution and equal efficiency between family and hired labour. His tests also support a linear procedure for aggregating male and female labour, subject to a relative wage adjustment factor.

Elizabeth Field (1988) provides an interesting test for the heterogeneity between free and slave labour in the cotton plantation economy of the ante-bellum southern United States. Using a translog production function with free and slave labour as distinct inputs, she finds that the Hicksian elasticity of complementarity (HEC) between free and slave labour was positive on both small and large plantations. These two labour inputs were *q*-complements, in that a higher application of one increased the marginal product of the other.[12] She also tests for the separability of these two labour inputs from the other inputs in the production function and finds evidence for weak separability. Thus free and slave labor can be combined into some aggregate index of labour input; but a simple linear combination, implying that they were perfect substitutes, is not supported by the data.

Pitt and Rosenzweig (1986) provide a novel indirect test of labour heterogeneity and the underlying assumption of the recursiveness of the production and consumption decisions of farm households. Using household level data from Indonesia that identified the incidence and severity of ill-health suffered by household members, they test whether the bouts of illness significantly affect the household's labour

---

[11] Le is effective labour, L^F is family labour and L^H is hired labour, with θ representing a constant difference in their productivity in effective units.

[12] The definition and interpretation of the Hicksian elasticity of complementarity is discussed more fully in Chapter VI of this thesis, together with the need to mind the *p*'s and *q*'s when talking about substitutes and complements, as reminded by Seidman (1989). The empirical tests for the heterogeneity and separability of family and hired labour reported in Chapter VI is exactly analogous to the specification and tests carried out by Field (1988) for free and slave labour.

supply as well as farm profits. They find that while illness significantly reduced the labour supply, farm profits are unaffected. This is an indirect test for labour homogeneity: market substitutes apparently can be found for the significant illness-induced reduction in the farm family's own labour supply on farm cultivation. They interpret such a result as proof that family and hired labour are homogeneous inputs and also as evidence for the recursive structure of the production and consumption/ labour supply components of the farm household model.

(b) labour supply

The labour supply decisions of farm households have not yet been as extensively analyzed as labour supply functions of workers in developed countries; and reliable estimates of the various elasticities of labour supply are few. But the available evidence (Singh, Squire and Strauss 1986c1, Bardhan 1979, Rosenzweig 1980) show that labor supply responses of farm households are not uniform across different asset class and farm size variables. The effects of changes in the wage rate and asset income on the labour supply of individuals varies particularly between landless/small farm cultivators working mainly on the off-farm wage market and larger farm cultivators who may allocate their total labour supply both to own farm cultivation and to off-farm wage employment.

One of the main methodological issues in modeling the labour supply behavior of farm households that devote some family labour to own farm cultivation is the identification of the wage rate to represent the true opportunity cost of family labour at the equilibrium position. There are three general approaches taken on this issue.

The first and most common approach has been to relate the opportunity cost of family labour in all households to the observed off-farm market wage rate, irrespective of whether a particular individual actually works off-farm or not. This assumption is consistent with a recursive structure of a farm household model with homogeneous labour, where the labour supply decision can be modeled separately from the production decisions. This is the approach adopted in Bardhan (1979) and

Rosenzweig (1980) which were some of the earliest examples of labour supply estimation for developing countries using large data sets. Assigning a common observed market wage rate as the opportunity cost of family labour is not always valid even under the assumption of homogeneous labour. (See the subsequent discussion in Chapter III). This assumption is even more suspect if family and hired labour are allowed to be heterogeneous inputs.

A second approach to labour supply estimation for farm households in a developing country setting has been implemented by Jacoby (1993) and Skoufias (1994), where the opportunity cost for family labour is equated to an estimated value of the marginal product of labour in own farm cultivation. This is an adaptation of the general methodology for estimating a structural labour supply equation for self-employed workers applied to a farm setting. It is analogous to the treatment of labour supply in the presence of progressive income taxes by "linearizing" the budget constraint as suggested by Hall (1973). In this method the estimation is done in a two-step procedure. In the first step, a farm-level production function is estimated, conditioning on the optimal levels of family labour supply that is observed. From the estimated parameters of the production function the marginal product of family labour can be derived for each sample household. These estimated marginal products are the appropriate shadow wage for valuing family labour and are also used to derive the "shadow" profit of the farm household from land ownership, which is included as non-labour income in the labour supply equation. In the second step, the labour-leisure equilibrium choice can be modeled as if the farm household were a pure consumer household which faced a parametrically given exogenous wage and non-labour income. [13]

The third approach has been to jointly estimate the production and labour supply components of the farm-household model in a non-recursive structure, allowing for the shadow wage rate for family labour to be endogenously determined. This

[13] Further details of this approach are given in Chapter IV since the estimation strategy used in this thesis is a simpler adaptation of the approach suggested by Jacoby (1993).

approach, however, is analytically cumbersome and requires complex estimation methods. Examples are Lopez (1984) and Newman and Gertler (1994).

The data set used in Bardhan (1979) is a large sample of nearly 5000 rural households in West Bengal from the Indian National Sample Survey. For the sample of male agricultural workers the wage coefficient is significantly positive. With the same model specification, Bardhan finds an insignificant wage response for male labour supply in the sample of own-farm cultivators. The wage elasticity for male agricultural workers was estimated to be between 0.2 to 0.3. The wage response was particularly insensitive for female workers in both types of households, with some indication of a backward bending labour supply curve for female labour supply. Other variables such as caste and marital status appear to be more important determinants of female labour supply.

Bardhan (1984) repeats the labour supply specification of his 1979 paper with another sample of households in West Bengal from a different round of the National Sample Survey. In this sample the wage effect is negative for male labour supply, which is opposite of the result obtained in the earlier paper. The wage effect is positive but insignificant for female labour supply. In both studies Bardhan develops a variety of wage measurements based on the reported market wage to accurately reflect the opportunity cost of family and market labour; but the wage sensitivity is quite low in almost all specifications.

Rosenzweig (1980), also using data from India and a labour supply equation specified with both male and female village level wage rates, finds that the supply of male labour is strongly backward bending for both landed and landless households. The labour supply curve for female workers is positively sloped with respect to the own wage, and the cross-wage effect is negative and significant.

Jacoby (1993) applies the sequential estimation method to peasant farms in the highland region of Peru where family based subsistence farming is concentrated. In the production function estimation Jacoby specifies three labour inputs - family male, family female and total hired. Jacoby tests for the separability of male and female family labour from the other inputs and finds the underlying production technology exhibits non-separability. Hence a translog function, with different interaction terms for these three distinct labour inputs is specified from which the marginal products of male and female family labour are estimated. This specification, however, results in negative values of the estimated marginal product of family labour, particularly for female labour, in nearly 20% of the sample. Jacoby is forced to drop these observations from the labour supply estimation step. For the subset of the sample with positive marginal products, the mean value of the marginal product of female labour is about 60% less than that of male family labour.

The labour supply estimation results in Jacoby (1993), using an instrumented version of the computed marginal products as the shadow wage rate for the equilibrium allocation of family labour, are reasonable. Uncompensated own-wage elasticities are positive for both male and female workers and the income elasticities are negative. This leads to significantly positive compensated own-wage elasticities, which is consistent with utility maximization. The cross-wage effects in both the male and female labour supply equations are positive (implying male and female leisure are complements in the household utility function); but these cross-wage effects are small relative to own-wage effects. Male own-wage elasticities are higher than for female labour supply, which is counter to the usual result of higher own-wage responsiveness of female labour supply, at least in developed country settings (Killingsworth and Heckman 1986). Jacoby indicates this result is mainly due to the definition of labour supply he has used, which includes domestic household work. Such a definition of female labour supply is likely to have a lower wage elasticity.

Skoufias (1994) has applied the same methodology used by Jacoby to the ICRISAT data for India, in a panel data framework with fixed effects. He uses a Cobb-Douglas production function with four different labour inputs, distinguished by gender and family and hired source. While the fit of the estimated production function is very reasonable (the R-sq. is 0.92 in the fixed effects model) the computed values of the marginal products of labour at the mean of the data provide very divergent results for the productivity of the four categories of labour. For instance, the marginal product of female family labour is estimated to be 4.7 times the marginal product of hired female labour. Similarly, the marginal product of female family labour is more than double the marginal product of male family labour. [14] There are no clear reasons to expect such large divergence in the marginal products by gender at the production optimum point. This result is likely to be a problem of the Cobb-Douglas specification with the four labour variables as distinct inputs.[15]

The labour supply results in Skoufias (1994) are also not fully satisfactory, especially for female family workers. Using the instrumented values of the estimated marginal products of family labour as the shadow wage rates, the uncompensated own wage effect is positive in the male labour supply equation, but the (female) cross-wage effect is negative, indicating male and female leisure are gross substitutes in this data set. The income effect is significantly negative for male labour supply, which is consistent with leisure being a normal good. Skoufias however obtains unusual results for the female labour supply equation - the gross uncompensated effect is negative, suggesting a backward bending labour supply curve. The explanation offered is the same as Jacoby's - that the labour supply variable includes all work activities including housework. But alternative estimates for labour supply excluding housework are not provided.

A backward bending labour supply relation can, of course, be a valid result for the gross uncompensated own-wage effect. But in Skoufias' results, because of the insignificant value of the negative income effect for female labour supply, even the compensated own-wage effect is negative. This is inconsistent with utility maximization. These anomalous results are likely to have been caused by the widely

[14] Skoufias (1994) does not directly report the estimated marginal products of the four labour categories in his paper. They can be computed from the estimated parameter values for the production function and the summary of the data reported in his Table 1. The marginal products (per hour), estimated at the mean of the data, are Rs. 0.87 for family male labour, Rs. 2.19 for family female, Rs. 0.27 for hired male and Rs. 0.47 for hired female categories.

[15] As reported above in part (a) of the literature review section on labour heterogeneity, Laufer (1985), using the same ICRISAT data set as Skoufias, reported a lower marginal product of female labour for all her crop-specific equations for the generalized quadratic production function. She confirms that her results are consistent with observed lower market wage rate for female labour.

divergent estimates of the marginal products of labour from his Cobb-Douglas production function specification with four distinct labour inputs.

Lambert and Magnac (1994) carry out only the first step estimation of the production function for a farm household model with distinct labour inputs; but they provide direct comparisons of the estimated marginal products and the market wage rates. Their results also indicate the problems inherent in deriving the marginal product of labour with many categories of labour defined in the production function. Using data for the Ivory Coast, they estimate a generalized linear production function and compare the estimates of the shadow price (marginal product) of labour with the market wage rates. Their production function specification distinguishes three different labour inputs - hired labour, family male and family females. Their estimates reveal large differences between the marginal product of these three labour inputs, which, while indicative of labour heterogeneity, are puzzling. For instance, in one set of estimates using instrumental variables, the marginal products of labour per day are estimated to be 157.2, 0.62 and 3.14 for hired labour, male family labour and female family labour, respectively.[16] These large differences remain unexplained and are even more puzzling given that the average market wage rate reported in their data is 5.76. Again, this is probably an indication of mis-specification of the production function with these three different labour input categories.

Given that Jacoby, Skoufias and Lambert and Magnac all find a very weak correlation between the estimated marginal products and observed market wage rates, Strauss and Duncan (1995) question the validity of using the estimated marginal products as proxies for the shadow wage rate, particularly if the estimation is based on survey data that does not have a detailed farm management component.

Newman and Gertler (1994) develop and apply to Peruvian data the more general joint methodology for estimating the labour supply behaviour of self-employed

[16] The units for the marginal products are not clearly reported since the output variable is an index. Presumably these are in hundreds of the Central African Franc (CFA), which is the unit in which the market wage rates are reported.

family members in a non-recursive framework. They specify three different equations: a direct specification of the marginal rate of substitution (MRS) between leisure and consumption to reflect household preferences, a market wage function for each individual in specific age-sex categories, and a marginal return to farm work function for individual by each age-sex categories.[17] This last equation is specified in terms of prices and quasi-fixed inputs and human capital related variables. It can be derived as the derivative of an implicit farm profit function with respect to farm labour for an individual in a specific age-sex group, conditional on the total application of family labour on the farm. This procedure allows them to model the heterogeneity in family labour across specific age-sex groupings both in the production and consumption components of the model.

On the consumption side, the MRS between leisure and consumption is, of course, not directly observable. But certain imputations can be made based on the first order conditions that equate the MRS to the marginal returns to labour. For individuals who report working in the off-farm labour market, the wage they receive will be equated to the MRS. For individuals who work only on the family farm, the shadow wage, which is equated to the MRS, can be computed from the marginal returns function. Since there is a dependence in both the underlying profit and utility function of one category of family labour on other categories, and since corner solutions for non-labor participation have to be explicitly modeled, this method of estimating the farm household model involves a great deal of econometric complexity, especially as household size and labour categories are increased. No other application of this approach appears to have been published to date, testifying to the great difficulty in implementing this methodology.

---

[17] Newman and Gertler's estimation procedure could actually be done with a separate marginal return to labour and MRS function for each household individual, conditional on the labour supply behaviour of other household members. But this will only increase the already stupendous amount of time taken to iteratively solve and estimate their model. The authors report that a single run of their model took all night on a 486 personal computer.

(c) Nepal-specific literature

The production component of farm household activities in Nepal have been studied in a growing number of empirical studies, based on production and profit function estimation strategies. But the specific question of heterogeneity between family and hired labour has not been directly addressed before. The labour supply and consumer demand system estimation for Nepalese rural households have been relatively neglected. When it comes to an integrated farm-household model estimation strategy, this author is aware of only one other study, Acharya (1987), which has adopted such a theoretically consistent framework. Even then, due to deficiencies in the data, the production and labour supply components in Acharya (1987) were not consistently linked in the actual estimation. [18]

Agricultural production function based studies at the household level in Nepal reveal a great deal of divergence in the regional patterns of production and input use both between the northern hill and southern lowland (*tarai*) region, as well as among the five development administrative regions which are defined on a east to west basis (NRB/ADB 1994). This variation indicates that in any large sample, with nationally representative data, regional fixed effects on productivity and labour use patterns are going to be important factors to consider.

Because of a higher man-land ratio, the northern hill area agriculture is labour intensive and several studies have estimated a very low marginal productivity of labour in hill agriculture - i.e. Belbase (1985). Similar direct estimates of the marginal product of labor for the *tarai* region agriculture are not available. Belbase and Grabowski (1985) and Grabowski and Belbase (1986) show divergences in technical efficiency of farms varying according to farm size; but they do not

---

[18] Acharya had to use a different data set to estimate the production function and labour supply components of the farm household estimation. In addition, her labour supply data is of limited quality since it does not have information on the actual hours or days of work.

discuss the underlying causes.[19] The issue of the work intensity of hired labour in particular, and the more general question of heterogeneity between different labour categories in Nepalese agriculture, have not been adequately addressed.

Mudhbary (1989) estimates an AIDS consumer demand system focussed only on the demand for various food categories. Labour supply/leisure demand is not considered. Several other studies look at labour supply in the limited sense of off-farm market employment, ignoring the labour supplied on own farm cultivation. An example is Rauniyar (1986) who provides labour supply estimation results for the off-farm work days only from a small household sample drawn from two districts in Nepal. He reports significant positive own wage coefficients for the labour supply behaviour of both men and women. As expected the incidence of off-farm labour supply is negatively related to farm size

The farm production component in Acharya (1987) is modeled with a quadratic profit function from which the first order conditions for the optimal factor demands provide the estimating equations. Acharya treats family labour as quasi-fixed inputs and estimates only the demand for hired labour conditional on the available stock of family workers. She finds that the demand for hired labour is negatively correlated to the own wage for both male and female labour. The cross-wage elasticities suggest male and female labour inputs are complements - an increase in the male wage rate also reduces the demand for female hired labour. She attributes this result to the strict gender based division of labour so that an increase in any wage reduces the demand for all types of labour. Within both gender categories, family labour is a substitute for hired labour. She estimates the hired labour demand system separately for two crops and four regional samples, and there is no clear uniformity in the relationship between the four labour categories. Also some of elasticity estimates are unreasonable - for example, the own wage elasticity for male hired labour demand is in excess of (plus) 5.0 (Acharya 1987: 153, Table 5.3).

---

[19] Moll (1990) has raised a strong objection to the findings reported in Grabowski and Belbase (1986), claiming their results of lower technical efficiency on large farms is due to the curve fitting nature of their estimated production function, and not due to a test of the hypothesis of lower efficiency on larger farms. See Grabowski (1990) for a response.

On the labour supply estimation component, Acharya estimates labour force participation and days of work equations separately for male and female household members. The dependant variable in the labour supply equation is not a scale variable that reflects the actual amount of time spent on specific work activities. The dependant variable was assembled from time allocation data that recorded only the frequency with which an individual engaged in specific activities within the observation period. The actual duration of each activity was not recorded. Even with such a limited proxy variable for labour supply, Acharya reports very reasonable results. Both the labour market participation decision and the frequency of work variables have positive wage effects and negative income effects.

The own-wage elasticity of labour supply was estimated to be 0.17 for males and slightly higher at 0.24 for female family members. The non-wage income effect, represented by the value of farm assets, is also significantly negative for both genders. These theoretically expected results, especially for female labour supply are significant, since in many other settings female work force participation and total working hours have been found to be unresponsive to wages and asset income (Ho 1980), or perhaps even having a backward bending relationship with respect to the wage rate (Bardhan 1979, Skoufias 1993).

Apart from wage and non-labour income, Acharya's results indicate that ethnic and caste grouping and demographic characteristics, such as family size and the number of adult workers per household also affect individual level labour supply decisions.

Cooke (1998) provides an example of a cross-sectional estimation of the demand for labour in the hill region agriculture of Nepal, using the observed market wage rate to value family labour. The main focus of Cooke's analysis is to determine how labour allocation of the family farm is affected by the scarcity of environmental goods - such as firewood, fodder and drinking water - that rural households have to collect with family labour, especially of women. She estimates a labour demand system that includes as regressors the wage rates for male and female labour, as well as the shadow price of the environmental goods, which are the wage-based valuation of the time taken to collect them. She assumes that male and female labour inputs are imperfect substitutes in production, but within each gender family and hired labour are perfect substitutes. With this specification she finds that the demand for labour is either insensitive to the human labour wage or increasing with the wage.[20] The latter result is theoretically inconsistent. Hence, the market wage may not accurately reflect the opportunity cost of family labour on the demand side as well.

The studies noted above provide a scattering of empirical work on rural households in Nepal. A complete farm household estimation strategy that also considers the question of labour heterogeneity between family and hired labour has not been carried out to date.

Another study related to the empirical work of this thesis is Hamal (1992) which presents the results of a profit function estimation for Nepalese agriculture, using aggregate national data for the 1961 to 1987 period. While Hamal's main significant finding is that total factor productivity growth in Nepalese agriculture, calculated from a growth accounting framework, has been negative during this period,[21] he also carried out a detailed empirical estimation of a translog cost function. He specified four main inputs: land, homogenous labour, bullock power and chemical fertilizers. He computes the Allen partial elasticity of substitution between these inputs for various time sub-periods. All of the estimated AES are positive - implying all pairs of inputs are *q*-substitutes. The key estimate of the AES between land and labour is equal to one in all the specific time periods of his estimates; and most of the other AES are also very close to one, indicating an underlying Cobb-Douglas primal technology.

[20] In Cooke's estimation results the coefficient on the human wage rate variable will of course be affected by the presence of the shadow price variables for the environmental goods. These shadow prices are just the valuation of the time taken to collect these goods, using the market wage rate to value an unit of time. So there is likely to be a strong correlation between these price variables. A separate labour demand equation without the other shadow wage variables was not reported. With respect to her main hypothesis, Cooke finds that labour input in farm cultivation is unaffected by the time taken to collect environment goods as reflected in her shadow prices for environmental goods.

[21] This result is consistent with the fact that large tracts of forested land, particularly in the *tarai* region, have been brought under cultivation without a commensurate increase in farm output.

2.3 Summary

While there is a large and growing literature on the estimation of farm household models for developing country settings, the specific issue of labour heterogeneity between family an hired labour has not been addressed in an integrated farm household model framework. There are a few studies that have looked at labour heterogeneity purely from a production function perspective. Production function estimates reveal a low degree of substitution between family and hired labour as well as different marginal effects in output; but the results are mixed and affected by the tendency to use limited functional form specifications. The specific sources of heterogeneity are not clearly specified.

Attempts to estimate a non-recursive model with many different types of labour category lead to puzzling results with very large discrepancies in the marginal product of the different labour categories. Specifications based on an aggregate labour composite, allowing for heterogeneity between the different components, are likely to do better

On the labour supply component, the approach most commonly used in the past (based on assigning the market wage rates of all farm households) produces mixed results. While more elaborate approaches to modeling the self-employment of family labour in peasant farming have been developed, these have not directly related the shadow wage rate applicable for family labour to the nature of the heterogeneity between family and hired labour. Nor do they take account of the fact that shadow wage rates should vary according to the labour hiring status of the farm household.

Alternative strategies of estimating labour heterogeneity and labour supply in a consistent manner have not been satisfactorily addressed in previous studies in general and much less so in the context of Nepalese agriculture. In the case of Nepal, even ignoring the labour heterogeneity issue, there has been a very limited focus on the empirical modeling of the decision-making framework of rural farm households.

CHAPTER III

A FARM-HOUSEHOLD MODELWITH HETEROGENEOUS LABOR INPUTS

3.1. Introduction

This Chapter presents the analytical structure of a farm household model allowing for family and hired labour to be either heterogeneous or homogeneous inputs in production. The model structure allows for different types of heterogeneity, and homogeneity of the labour inputs occurs as a special case of the general model. The focus is not on the specific source of this heterogeneity, but the manner in which it affects the analytical structure of the basic farm household model. A key objective of this Chapter is to show under what conditions the recursive property of farm household models is maintained, even with heterogeneous labour inputs; and to derive the implications of this structure for estimating the labour supply component of the model.

Section 3.2 presents a model where family and hired labour are treated as completely separate inputs, which is the most general form of labour heterogeneity. The next section discusses the various difficulties in specifying and in estimating a farm household model of this general form. Section 3.4 presents an alternative model with a nested production structure where in the first stage family and hired labour are combined to create an aggregate or composite labour unit which then enters into the farm production function in the second stage. Section 3.5 derives the labour supply implications of the nested production structure with heterogeneous labour. One important advantage of a farm household model structure with a nested aggregate labour input is that the "shadow wage rate" for family labour can be related to the observed market wage rate for hired labour, when both family and hired labour are simultaneously used on the family farm. Such a relationship helps in identifying the correct effective wage rate to be used in the estimation of labour supply functions for family labour even when it is applied solely to the family farm.

The structure of the farm household model presented in this Chapter is focussed on the treatment of family and hired labour as heterogeneous inputs. Many other important and analytically interesting issues that can be incorporated into farm household models (i.e., home goods production, marketable surplus of farm production) are not incorporated into the models presented in this Chapter.[1] It also adopts the convention of treating a multi-person farm household as a single decision making unit which maximizes household level utility, ignoring decision making based on individual utility maximization and bargaining among the household members.[2]

3.2. Analytical Structure of a Model with Heterogeneous Labour

Let U represent an aggregate household utility function defined over a composite consumption good C and leisure $\mathcal{L}$. Let T be the total labour time endowment of the household that has three components: own-farm labour use (F), market wage employment (M) and leisure ($\mathcal{L}$). The inputs in the farm production process are land area cultivated (A), family labour (F) and hired labour (H); and the two types of labour are initially treated as distinct inputs. For simplicity assume there is a single (or composite) farm output, and other variable inputs in the production process are ignored because they have little bearing on the treatment of the heterogeneity between family and hired labour inputs[3]. The farm household also has non-labour endowment income of E.

[1] This approach also ignores the Beckerian tradition of treating household production of consumption goods based on inputs of purchased goods and family labour time. Becker's (1965) new household economics is a separate development that is mainly applied to developed country settings.

[2] See Chiappori (1988, 1997) for household models with individual bargaining.

[3] The main analytical relevance of explicitly modeling other variable inputs is that the presence of other variable inputs may affect the marginal product of the two types of labour in different ways. If the ratio of the marginal product of family labour to that of hired labour is sensitive to the level of other inputs, the production function is not *separable* in the two labour inputs (Chambers, 1986: 43). This means the two labour inputs cannot be consistently aggregated into a composite labour input. In a particular setting whether family and hired labour are separable inputs is an empirical question; and this is addressed in Chapter VI. Note this use of the term "separable" as a property of the production function is different from the "separable" (recursive) property of farm household models.

Let **p** represent the price of farm output and $\mathbf{p}_c$ (normalized to 1) the price of the composite consumption good. Let $\mathbf{w}^h$ be the wage rate at which labour can be hired in and **w** the net wage rate which family labour receives when it is supplied in the hired labour market. (It is common to find $\mathbf{w}^h > \mathbf{w}$ because it is likely there will be some costs associated with working away from one's home; but no restrictions on these wage rates are imposed *a priori*). For simplicity assume that the net wage rate received for all off-farm work is uniform - i.e. **w** applies to all alternative sources of employment for family labour outside of its own-farm cultivation, whether it be agricultural work on big farms or non-agricultural work within or outside the rural areas. It is assumed that labour markets clear and households do not face binding quantity constraints on either their labour demand or labour supply. The normal seasonal fluctuation in agricultural wages is also ignored (although the analytical framework of the model with a uniform wage will be valid in each of the specific time periods during which the seasonal wage is assumed fixed).

The household's maximization problem can then be set up as follows :

(3.1) $\max \quad U(C, \mathcal{L})$

subject to

(3.2) $C = pQ + wM - w^hH + E$

(3.3) $Q = f(F, H, A)$

(3.4) $T = \mathcal{L} + F + M$

(3.5) $F \geq 0;\; M \geq 0;\; H \geq 0$

The constraints 3.2 to 3.4 deal, respectively, with the household cash income constraint for the purchase of the consumption good (C), the farm production technology constraint, and the household total time endowment constraint. The last restriction (3.5) imposes non-negativity constraints on the labour categories, allowing households the option to set M, F and H to zero.

It is analytically convenient to combine the three budget constraints represented by Eq. 3.2 to 3.4 into a single constraint referred to as the "full income" constraint. This will be of the form[4]

$$(3.2n) \qquad C + w\mathcal{L} = \{pQ - wF - w^hH\} + wT + E$$

$$= \Pi + wT + E$$

where Π represents the (short run) profits from farm production derived by deducting the cost of family and hired labour from gross output, with family labour valued at the market wage rate w. The total net returns to the farm household from its own farm production is broken down into two components: Π, which measures the net profit or returns from the ownership of the farm land, and wF, which represents the wage labour income from own-account work on the family farm.

The LHS of Eq. 3.2n represents the total expenditures of the household on consumption of goods and leisure. The RHS represents the *full income* of the farm household that has three components: farm profits (Π), the value of the total time endowment of the household (wT) evaluated at the wage rate it receives when working off the farm, and the non-labour endowment income, E.

To solve the household maximization problem specified above, substitute the constraint (3.4) directly into the utility function to set up the following Lagrangian:

$$(3.6) \quad \text{Max } \pounds = U[C, (T - F - M)] +$$

$$\lambda [pQ(F,H,A) - w^hH + wM + E - C] + \mu_1 M + \mu_2 H + \mu_3 F$$

where μ_1, μ_2 and μ_3 are the Lagrangian multipliers associated with the non-negativity constraints on M, H and F which satisfy the following complementary slackness conditions :

$$(3.7a) \qquad \mu_1 * M = 0 \implies \mu_1 = 0 \text{ if } M > 0$$

$$(3.7b) \qquad \mu_2 * H = 0 \implies \mu_2 = 0 \text{ if } H > 0$$

$$(3.7c) \qquad \mu_3 * F = 0 \implies \mu_3 = 0 \text{ if } F > 0$$

[4] Equation 3.2n is derived by adding $w\mathcal{L}$ to both sides of Eq. 3.2 and substituting out for $\mathcal{L}$ on the left hand side using Eq. 3.4.

The choice variables for the household are C, M, F, H for which the first order conditions, respectively, are [5]

$$(3.8) \qquad U_C = \lambda$$

$$(3.9) \qquad U_{\mathcal{L}} = \lambda w + \mu_1$$

$$(3.10) \qquad U_{\mathcal{L}} = \lambda (p \frac{\partial Q}{\partial F}) + \mu_3$$

$$(3.11) \qquad \lambda (p \frac{\partial Q}{\partial H} - w^h) + \mu_2 = 0$$

From (3.8) and (3.9) it follows

$$(3.12) \qquad \frac{U_{\mathcal{L}}}{U_C} = w + \frac{\mu_1}{\lambda}$$

The left-hand side of Eq. 3.12 represents the marginal rate of indifferent substitution of leisure for consumption ($MRS_{\mathcal{L}C}$) in the household's utility function. It measures the household's subjective valuation of leisure foregone in terms of the consumption numeraire when the household supplies an extra unit of labour.

Let L represent the total labour supply of the farm-household ($L= F + M = T - \mathcal{L}$). Since

$$(3.13) \qquad \frac{U_{\mathcal{L}}}{U_C} = - \frac{U_L}{U_C}$$

where U_L is the marginal dis-utility of labour, the left hand side of Eq. 3.12 measures the subjective value placed on a marginal unit of household labour. Sen (1966) refers to this as the "the real cost of family labour" whereas Nakajima (1986) calls it the "marginal (subjective) valuation of family labour".

From (3.8) and (3.10) it follows

$$(3.14) \qquad \frac{U_{\mathcal{L}}}{U_C} = p \frac{\partial Q}{\partial F} + \frac{\mu_3}{\lambda}$$

---

5 Following conventional notation $U_C = \partial U / \partial C$ and $U_{\mathcal{L}} = \partial U / \partial \mathcal{L}$.

Equation 3.11 can be re-written as:

$$(3.11a) \qquad p\frac{\partial Q}{\partial H} = w^{h} - \frac{\mu_2}{\lambda}$$

The farm household equilibrium with respect to labour supply and labour demand are then given by Equations 3.11a above and 3.12 and 3.14 (repeated below).

$$(3.12) \qquad \frac{U_{\ell}}{U_C} = w + \frac{\mu_1}{\lambda}$$

$$(3.14) \qquad \frac{U_{\ell}}{U_C} = p\frac{\partial Q}{\partial F} + \frac{\mu_3}{\lambda}$$

These three Equations above represent the household labour supply and labour demand equilibrium for all possible combinations of the three labour categories (M, F and H). From Eq. 3.11a it is clear that if any hired labour is used (i.e. when $H > 0$, implying $\mu_2 = 0$), the value of the marginal product of hired labour is set equal to the wage rate paid out (w^h). And if no hired labour is used, it must be true that the value of the marginal product of the first unit of hired labour is less than w^h (since both λ and μ_2 are non-negative).

Similarly, from Eq. 3.14 if no family labour supply is used for own farm cultivation ($F = 0$, implying $\mu_3 \geq 0$) then the household's valuation of its leisure must be no less than the marginal product for the first unit of family labour applied to the farm.

For the household that supplies some of its labour on own-farm cultivation and some labour also to the off-farm labour market at wage w, the household equilibrium is characterized by:

$$(3.15) \qquad \frac{U_{\ell}}{U_C} = w = p\frac{\partial Q}{\partial F}$$

This is the standard first order condition which specifies that when family labour has two different uses - working on the family farm (F) and working in the off-farm labour market (M), the returns to labour in both activities are equalized at the margin. Hence family labour is applied in own farm production until the point where the value of marginal product is equated to the market wage rate for off-farm work.

For the optimal family labour supply not to involve any market work (ie, for when M = 0), the equilibrium condition is:

$$(3.16) \qquad \frac{U_{\ell}}{U_C} = w^* = p\frac{\partial Q}{\partial F}$$

where $$w^* = w + \frac{\mu_1}{\lambda}$$

and $$w^* \geq w, \quad \text{since } \lambda > 0 \text{ and } \mu_1 \geq 0 \quad \text{if } M = 0.$$

The w* represents a "shadow-wage rate" that applies to the subjective equilibrium of the labour-leisure choice faced by the farm household that only works on its own farm (Nakajima 1986). If a farm household does not supply any family labour to the off-farm labour market at a net wage of w, then the marginal returns to applying family labour to own farm cultivation must be no less than the market wage rate w.[6]

When M, F and H are all positive, such that all of the μ 's are zero, the household equilibrium is characterized by the conventional optimality conditions. These equate the marginal products of the two types of labour inputs to their respective wage rates, and also the real cost of family labour to the market wage rate. These optimality conditions are given by Eq. 3.15 and Eq. 3.11a with $\mu 2 = 0$.

Diagrammatic illustrations

The above equilibrium conditions for the allocation of family labour are illustrated in Figure 3.1 for the case where some family labour is applied on the farm and also is hired out (F > 0, M > 0). Figure 3.2 illustrates the case where all of the family labour supply is devoted to own farm cultivation (M = 0).

In Panel A of Figure 3.1 the vertical axis measures the money value of consumption (C), say in Rupees, while the horizontal axis measures family labour/ leisure time, say in days. Since OT represents the total time endowment of family labour, the

[6] This result holds under the assumption there are no quantity constraints on the days of labour supplied and no fixed costs to working on the off-farm labour market.

distance to the right from point O measures family labour work days while the distance to the left from point T measures leisure days. The bold lines I_1 and I_2 represent the farm-household's indifference curves between consumption and leisure. These indifference curves as drawn have the conventional property of being downward sloping and convex from below, given that in the $\{C, \mathcal{L}\}$ space the origin is represented by point T.[7] The upward sloping curved line OGQ represents the farm production curve as family labour increases, measured in terms of the consumption good. Since $\mathbf{p}_c$ is normalized to one, OGQ represents the value of farm production as family labour input is varied, holding all other inputs (including hired labour) fixed.

The farm-household equilibrium in Panel A is represented by the points of tangency of the line DGK with the farm production curve at the point G and with the indifference curve at the point J. The line DGK has the slope **w,** which is the off-farm market wage rate for family labour. The tangency at point G represents the optimal family labour applied to own farm cultivation in the amount of F, while the tangency at point J represents the total labour supply of the farm household, which includes F and the amount M of work in the off-farm labour market at wage w.

Panel B in Figure 3.1 gives an equivalent representation of the labour-leisure equilibrium depicted in Panel A. The vertical axis of Panel B is in monetary units while the horizontal axis is labour days worked as in Panel A. The connection

---

[7] Along the indifference curves I_1 and I_2 as leisure increases the slope of the indifference curve is falling, indicating a decreasing marginal rate of substitution between consumption and leisure as more leisure is taken. More formally sufficient conditions to ensure that indifference curves are downward sloping in the $\{C, \mathcal{L}\}$space and convex to the origin are

(N3.7.1) $$\frac{\partial U}{\partial C} = U_C > 0 \quad ; \quad \frac{\partial U}{\partial \mathcal{L}} = U_{\mathcal{L}} > 0$$

(N3.7.2) $$\frac{\partial (U_{\mathcal{L}}/U_C)}{\partial C} = > 0 \ ; \ \frac{\partial (U_{\mathcal{L}}/U_C)}{\partial \mathcal{L}} = < 0$$

The implications of these restrictions are fully drawn out in Nakajima (1986:11 -14). Although Nakajima works directly with a utility function specified in terms of C and L (labour work days) instead of C and $\mathcal{L}$ (leisure days) as in the above, the restrictions above are exactly equivalent to the restrictions 2.2, 2.11 and 2.12 imposed by Nakajima to obtain well behaved indifference curves which slope upwards in the {C, L} space and which are convex from below.

between Panel A and Panel B is that a measurement represented by a slope in Panel A becomes a vertical distance in Panel B. Hence, the YY curve in Panel B is the value of the marginal product of family labour which traces out the slope of the OGQ curve from Panel A at the equivalent level of family labour input.[8] Similarly, the VV curve measures the marginal rate of substitution between leisure and consumption (U_{ℓ}/U_{C}) which is the (positive) value of the slope of the indifference curve between leisure and consumption) at different levels of family labour work-days. As drawn in Panel B of Figures 3.1 and 3.2, the VV curve is upward sloping in its entire range.[9] This need not be the case always and the VV curve can have flat sections at low levels of labour supply.[10]

In Panel B the solution relating to the optimal use of family labour in own farm cultivation is given by point G' - the equality of the market wage rate and the marginal value product of family labour (corresponding to the first equality in Eq. 3.15). The solution with respect to the optimal total labour supply of the farm household is given by point J' - the equality of the wage rate and the marginal rate of substitution of leisure for consumption which represents the marginal subjective valuation of family labour. (Point J' corresponds to the second equality in Eq. 3.15).

---

[8] The YY curve is downwards sloping in its entire range under the assumption of a regular production function with declining marginal productivity of labour (and of all other) inputs.

[9] The VV curve sloping upward in its entire range implies that the $MRS_{\ell C}$ is continuously increasing as work days (L) increases or leisure (ℓ) decreases. Following Nakajima (1986:26) the necessary and sufficient conditions for VV to be upward sloping are :

$$\text{(N3.9.1)} \quad Z = \frac{U_{\ell}}{U_C} > 0; \text{ and } \quad \frac{\partial Z}{\partial C} * p \frac{\partial Q}{\partial F} - \frac{\partial Z}{\partial \ell} > 0$$

These conditions are satisfied under the restrictions on the utility function specified by N3.7.1 and N3.7.2 of Footnote 3 above in this Chapter (assuming that the marginal product of family labour -- $\partial Q/\partial F$ -- is always non-negative).

[10] It is quite feasible that VV be horizontal at low levels of labour and consumption. For instance if one postulates a minimum subsistence consumption level C_0, and since low levels of labour input will lead to low levels of C (ignoring non-labour income), the MRS_{LC} may remain unchanged until the labour days worked is sufficient to attain the C_0 level of consumption. Graphically this means the indifference curves I1 and I2 when transferred to the region below C_0 would be upward sloping straight parallel lines with a constant slope in the {C, L} space. Consequently in Panel B, there would a section of the VV curve being horizontal at low levels of labour input. See Nakajima (1986 pp. 19-20). Horizontal sections of the VV curve are of analytical interest since it gives rise to "surplus labour" in the sense defined by Sen (1966).

Panel B in Figure 3.1 also clearly illustrates the separable property of the farm household equilibrium when labour is supplied on the off-farm hired labour market ($M > 0$). The optimal allocation of family labour for own farm cultivation represented by point G' is completely independent of the household's preferences with respect to leisure and consumption. Such preferences, which affect the shape of the indifference curves and hence the position and slope of the VV curve, determine only the total labour supply equilibrium position of E'. The VV curve has no effect on the position of D' which is determined solely by the exogenously given wage rate and the parameters of the farm production function which underlie the YY curve.

The equilibrium conditions for the labour allocation of the farm household when all family labour supply is devoted to the own farm ($M = 0$) is illustrated in Figure. 3.2. In Panel A the optimum allocation is given by the point G where the indifference curve I_1 is tangent to the value of farm production curve OGQ. The total labour supply of the farm household is OU days and all of it is supplied to the own-farm. At G the slope of the line of tangency, DN denoted as w^*, is higher than the slope of the wage line DK, given by w. When family labour is allocated to own farm cultivation only, the returns to own farm cultivation must be greater than the returns to working in the off-farm labour market at the going wage rate (as indicated by the first order condition of Eq. 3.16 with $\mu_1 > 0$). The value of w^* can be interpreted as "shadow" wage rate at which a hypothetical market for family labour is in equilibrium. A market wage rate of w^* would equate the demand for family labour in own farm cultivation to the supply of family labour offered at w^*. The equilibrium points G (and G' in Panel B) involve the farm household's subjective valuation of its labour which is determined jointly by production technology and preferences and is independent of the observed market wage rate, w.

The farm household model with $M = 0$ is no longer separable into a producer and consumer equilibrium. The optimal input of labour on the family farm (which consists solely of family labour) is now affected by the household's labour-leisure preferences. Alternatively, the profit maximizing behaviour of a firm cannot be separated out from the utility maximization behaviour of the consumer household.

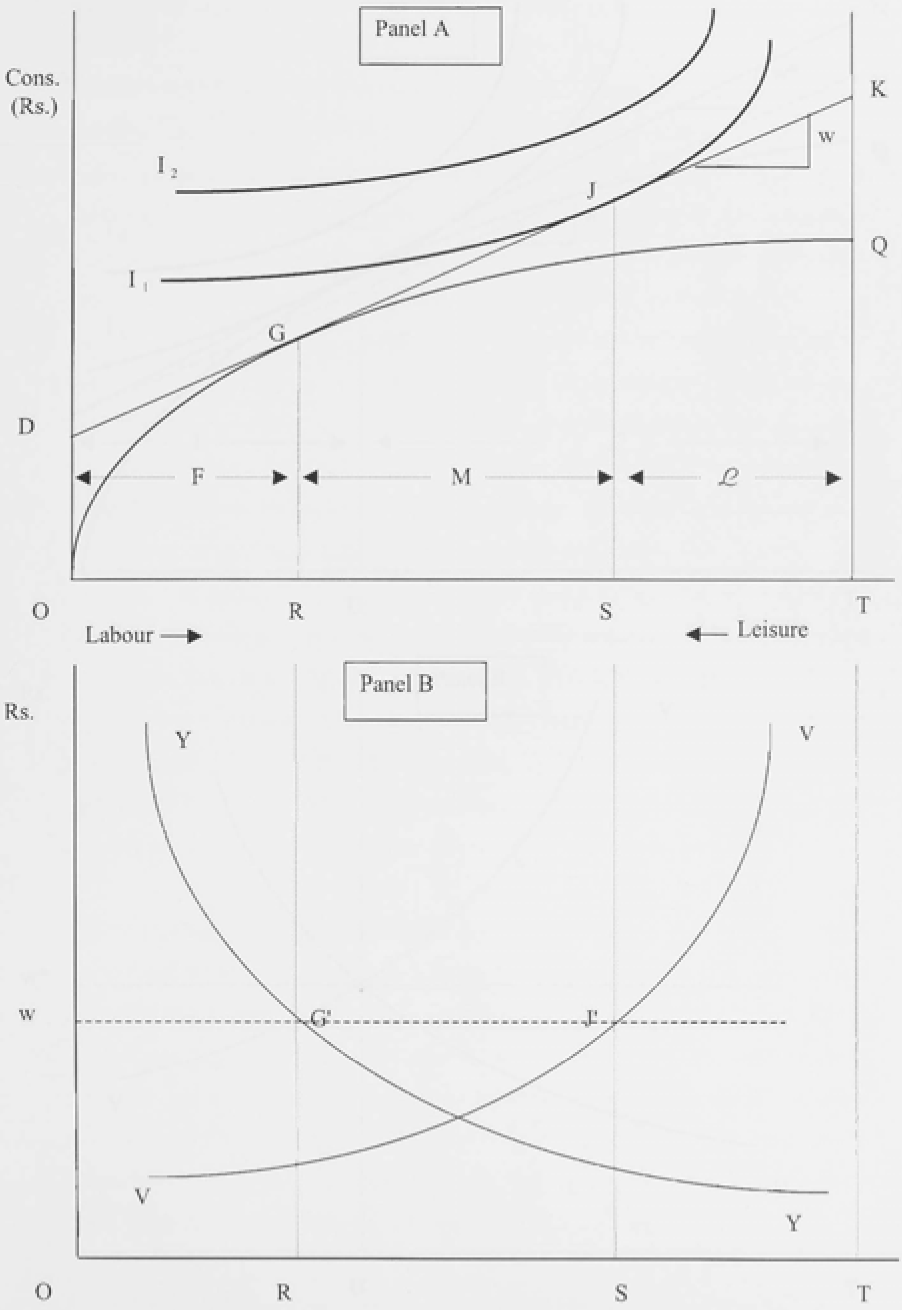


Figure 3.1: Farm Household Equilibrium with Family Labour Hired Out

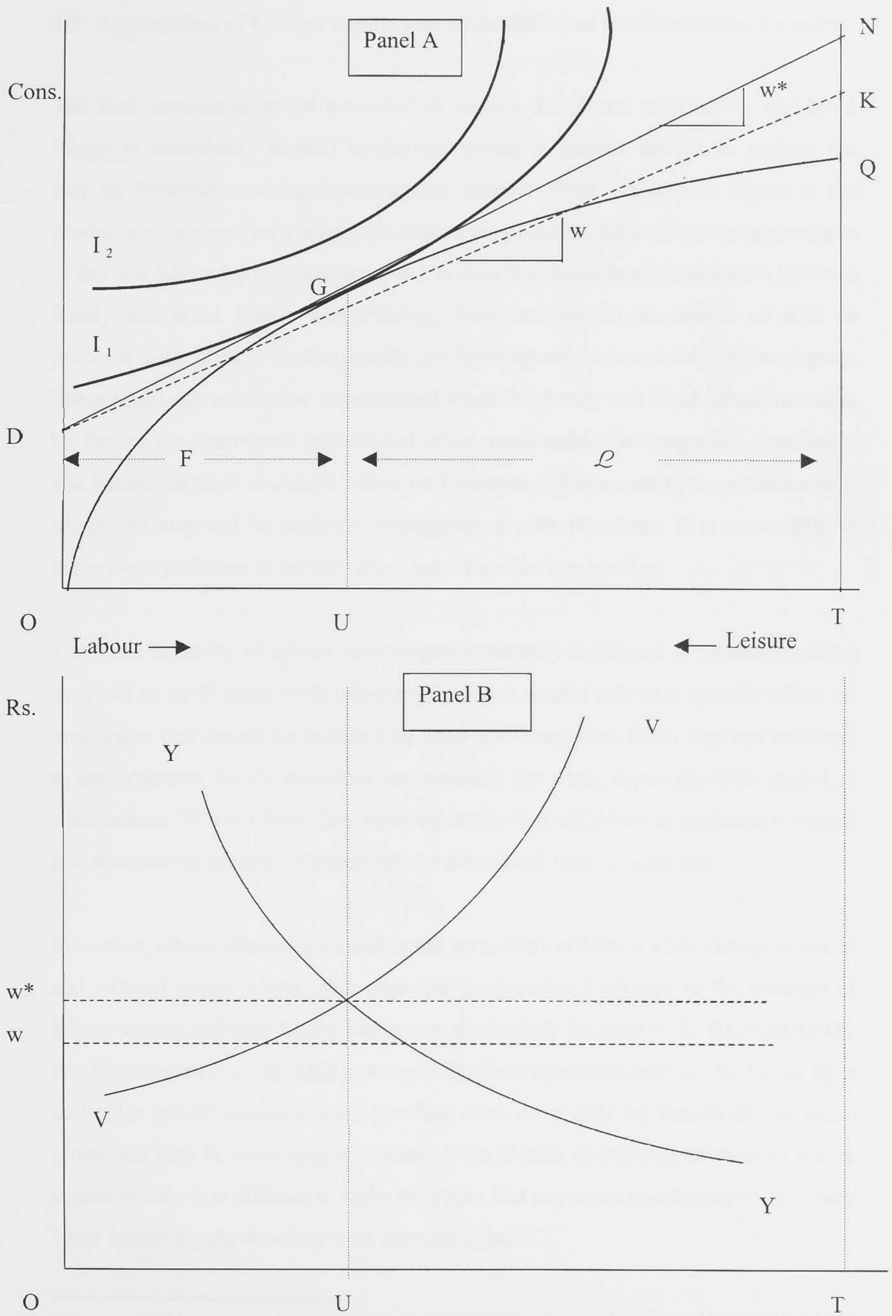


Figure 3.2: Farm Household Equilibrium with No Family Labour Hired

3.3 Aggregation of Labour Inputs and Separability of the Production Function

The farm household model presented in Section 3.2 above, with family and hired labour as completely distinct production inputs, is general enough to capture the various dimensions of any heterogeneity between them. As distinct inputs in the production function, no *a priori* restrictions are placed on how the marginal products of the two labour inputs are determined, nor on the elasticity of substitution between family and hired labour. Nevertheless, there are several theoretical as well as practical difficulties in treating family and hired labour as completely distinct inputs. These problems are further compounded when the family and hired labour can each be further dis-aggregated into several other meaningful sub-categories -e.g., male and female, or adult and child within each category. While each of these labour sub-categories may not be perfectly homogeneous with all others, it is reasonable to expect many of them to be very close substitutes for one another.

The wide diversity of labour sub-categories actually employed in peasant farming does not in itself mean each sub-category has a special role or a specific effect on production that cannot be matched by other sub-categories. Farm families will tend to use whatever family resources are available for work, especially with regard to child labour (White 1994). This does not imply that child labour performs a special role that cannot be done as effectively by other adult family members.

Moreover, labour allocation in traditional agriculture reflects a wide variety of social and cultural norms which often give rise to specialized patterns of the division of labour among different labour categories, particularly by gender (K. Bardhan 1993). For instance, in a given setting one specific farm operation may be done only by a particular labour category (e.g., weeding work done only by female labour while ploughing may be done only by males). Even if such an extreme division of labour occurs widely it is difficult to make the claim that any one sub-category of the many types labour inputs observed is an essential input.[11]

[11] An essential input must be used in order to have positive levels of output. If $F(x_1, x_2, \ldots, x_n) = 0$ whenever $x_i = 0$, then x_i is an essential input (Chambers 1988: 9).

In specific village settings, while there may be general conventions about the sexual division of labour and other labour categories, there often are sufficient exceptions to the rule to belie any claims for the *sui generis* nature of specific labour categories.

The main analytical problem with specifying a farm household model with many different labour categories is that the marginal product schedule for a specific type of labour (such as the YY curve in Fig. 3.1) can only be defined if the levels of the other labour inputs are taken as given. In an n input production function, while the marginal product of a particular input is affected by the levels of all n inputs, the relationship is specifically sensitive to the levels of other inputs which are very close substitutes (or very strong complements). With several different labour categories which are close substitutes for each other and none of which is by itself an essential input, there is a wide variety of ways of generating the same level of an aggregate or composite labour input. The marginal product schedule for any one labour input category will depend not only on the substitution possibilities between that particular labour input and all other remaining labour categories, but also on the substitution possibilities among each of the remaining labour categories.[12]

While it is technically feasible to draw a marginal product schedule for any one labour category, holding the levels of all other labour and non-labour inputs fixed, it is difficult to provide an economic explanation for why the other labour inputs are held constant at any particular level when there is a high degree of substitution among several labour categories. The marginal product schedule for family labour holding hired labour fixed at, say, 10 units will be very different from the one that holds hired labour fixed at 200 units. In order to make sense of the marginal product schedule for family labour around a specific equilibrium position it will be necessary to explain what is the optimal level of hired labour that goes with that particular equilibrium allocation of family labour. When family and hired labour are close substitutes for each other the relative distribution of the total labour inputs between

[12] In a n input production function with k labour categories, there are $k*(k-1)/2$ distinct partial elasticities of substitution among the labour inputs themselves. When only a few of these distinct partial elasticities are high, it becomes difficult to derive a marginal product schedule based on only one specific labour category without aggregating in some way all the other remaining categories.

family and hired labour may not be properly defined. Many different combinations of family and hired labour allocations could satisfy a given optimal level of total labour input.[13]

A related problem is that with many distinct labour inputs, it is often difficult to rationalize the observed division of labour among various labour categories with the optimal conditions of the equality of the ratio of the marginal products with the ratio of the wage rates. The wage rates for different types of labour are often specified in fixed proportions. For instance, the female adult wage rate may be 80% of the male adult wage rate in a wide variety of settings where it would be difficult to justify that the underlying productivity differences between male and female labour should remain at a constant 80% wherever this fixed wage gap is observed. In such a setting, a farm household model that had male and female labour as separate inputs would have difficulties in relating the observed division of labour by gender to the optimality conditions of the equality of the ratio of marginal products to the ratio of the wage rates. With a fixed wage gap and close substitutability, one would expect such models to predict a much higher incidence of complete specialization - using only male or only female labour - than what occurs in reality.

There also are practical difficulties in empirical estimation of production functions with many categories of labour as distinct inputs. The parameter space expands rapidly with many different labour categories as distinct inputs, especially when using flexible functional forms. There are special econometric problems when many observations in a data set may have zero usage of some specific labour categories. Chapter VI will show that the production function parameter estimates differ substantially between specifications that use distinct labour input categories and an aggregate composite category, even when using flexible functional forms.

[13] In the conventional farm household model where family and hired labour are treated as homogeneous inputs, the model determines only the total labour demand and labour supply. Any combination of family and hired labour on the demand side and any matching combination of own farm work and off-farm work on the labour supply side can satisfy these aggregate levels. Such indeterminacy of the individual components of F, H and M arise as the degree of substitution between family and hired labour increases when family and hired labour are treated as distinct inputs.

It is meaningful then to seek to combine all the different labour sub-categories into an aggregate or effective labour input, while still allowing for heterogeneity between different labour categories. However, it is possible to create a consistent labour aggregate only under special separability properties of the underlying production function (Berndt and Christensen 1974). Separability of the various labour-input categories from other non-labour inputs in the production function implies that the marginal products of all the labour input categories are affected in a uniform manner by the levels of the non-labour inputs.

Formally, in a n-input production function given by

$$(3.17) \qquad y = f(x_1, x_2, ..., x_n)$$

if the subset of labour input categories is indicated by the inputs x_1 to x_k, the production function is "separable" in the labour sub-categories if the production function can be written in an equivalent form as

$$(3.18) \qquad y = f\{g(x_1, x_2, ..., x_k), x_{k+1}, ..., x_n\}$$

where the g(.) function itself is quasi-concave and strictly monotonic.[14]

The g(.) function in itself has all the properties of a regular production function. It can be interpreted as a *micro* production function in which the output is the aggregate labour variable and the inputs are the different categories of labour.

Separability implies that the marginal rates of substitution between pairs of factors in the separated group are independent of the levels of factors outside that group.[15] Whether family and hired labour are separable from other inputs in a particular setting is a question that can be empirically verified with farm level data. If, in a given setting, the separability of the labour inputs can be established, then all the

[14] Goldman and Uzawa (1964).

[15] The formal representations of the separable property of a production function are discussed in more detail in Chapter VI. Note that "separable" as a property of the production function is different from the "separable" (recursive) property of farm household models. To avoid confusion henceforth, the term "recursive" will be used to refer to the property of farm household models while the term "separable" will refer only to the property of production functions.

different categories of the labour inputs can be aggregated. This leads to some simplifications in the structure of the farm household model structure presented in Section 3.2 above, and consequently the separability property greatly facilitates the subsequent empirical estimation even in the presence of labour heterogeneity.

The main theoretical implication of the separability of the labour inputs is that the ratios of the marginal product of family and hired labour (or alternatively, the marginal rate of substitution between family and hired labour) depend only on the levels of the labour inputs. Hence, optimal decisions about the allocation of family and hired labour can be made independently of the choices about other non-labour inputs. This gives rise to a sequential decision making (Chambers 1988:45). With separability the optimal levels of the labour inputs are chosen only with respect to the relative price ratios (wage rates) of the different labour categories. Hence there is now a clear relationship between the shadow wage applicable to family labour and the market wage rate for hired labour on farms that use both family and hired labour. This relationship is derived in the next section.

3.4 A Farm Household Model with a Heterogeneous Labour Composite

Assuming a separable structure of the labour inputs in the farm production function, the farm-household's maximization problem can then be set up in an identical manner to Equation 3.1 to 3.5 with the production function constraint (Eq. 3.3) written in the separable format as

$$Q = f(g(F, H), A) \tag{3.3a}$$

where $Le = g(F, H)$ defines an aggregate composite index of total labour input, which can be treated as units of effective labour created out of the observed levels of family and hired labour applied in farm production.

The choice variables for the household utility maximization problem are the same as in Section 3.2; and hence the first order conditions are identical to Equations 3.8 to 3.11. With the two types of labour inputs being separable in the production function, an additional property of the farm household equilibrium is that the ratios of the marginal products of family and hired labour is a function solely of the levels of family and hired labour inputs. That is,

$$(3.19) \qquad \frac{\partial Q / \partial H}{\partial Q / \partial F} = \frac{\partial g / \partial H}{\partial g / \partial F} = \mathrm{h(F, H)}$$

At the optimal values of family and hired labour input (indicated as F* and H*), let this ratio of the marginal products be given by θ*. That is

$$(3.20) \qquad \mathrm{h}\,(F^*, H^*) = \theta^*$$

This additional condition derived from the separability of the labour inputs provides a way to relate the wage rates for the two types of labour, using the first order condition that the marginal product of hired labour will be set equal to w^h, the hired labour wage rate.

From the first order conditions for the farm household equilibrium, using Eq. 3.12 and Eq. 3.14, after substituting out for U_{ℓ}/U_C, gives

$$(3.21) \qquad p\frac{\delta Q}{\delta F} = w + \frac{\mu_1}{\lambda} - \frac{\mu_3}{\lambda}; \text{where } \mu_1 = 0 \text{ if } M > 0; \quad \mu_3 = 0 \text{ if } F > 0.$$

From Eq. 3.11a, one gets

$$(3.22) \qquad p\frac{\partial Q}{\partial H} = w^h - \frac{\mu_2}{\lambda}; \quad \text{where } \mu_2 = 0 \text{ if } H > 0.$$

hence, the ratio of the marginal products can be expressed as

$$(3.23) \qquad \frac{\partial Q / \partial H}{\partial Q / \partial F} = \frac{w^h - \frac{\mu_2}{\lambda}}{w + \frac{\mu_1}{\lambda} - \frac{\mu_3}{\lambda}} = \theta^* = \mathrm{h}(F^*, H^*)$$

keeping in mind the various complementarity slackness conditions on μ_1, μ_2 and μ_3.

Equation 3.23 is the basic optimality condition that must be satisfied for all possible labour allocations involving various combinations of F, M and H. Equation 3.23 is particularly useful in the case when family labour does not work at all on the off-farm market at wage w (i.e. when M = 0) because in such situations the shadow wage rate for family labour can still be recovered from the observed levels of w^h.[16]

Consider first the case of a big farm household that must rely on hired labour to some extent (i.e. where $\mu_2 = 0$ because H > 0). For simplicity assume initially that $\theta^* = 1$ at the particular allocations of F and H that are observed. Then the shadow wage rate at the margin for valuing family labour, w*, is given by

$$(3.22) \qquad w^* = w + \frac{\mu_1}{\lambda} = \frac{w^h}{\theta^*} = w^h$$

Since one unit of family labour substitutes at the margin for one unit of hired labour which costs w^h, the relevant shadow price for family labour must also be w^h as long as hired labour is employed on the farm and $w^h > w$. In this situation family labour is never supplied on the off-farm market employment at wage w when it could be used to replace a unit of hired labour which costs w^h.

If instead $w > w^h$, then it would lead to complete specialization of the farm household labour allocation: all of its labour supply will be on market employment at wage w, and all of the labour input on the family farm will be hired labour paid a wage w^h per unit. Only in the case where $w = w^h$ would it be optimal for the big farm household to supply labour to the off farm labour market employment as well as on its own farm. This is readily evident from Eq. 3.23 since for both F and M to be positive requires $\mu_1 = \mu_2 = 0$; hence if $\theta^*=1$ it must be that $w = w^h$.

---

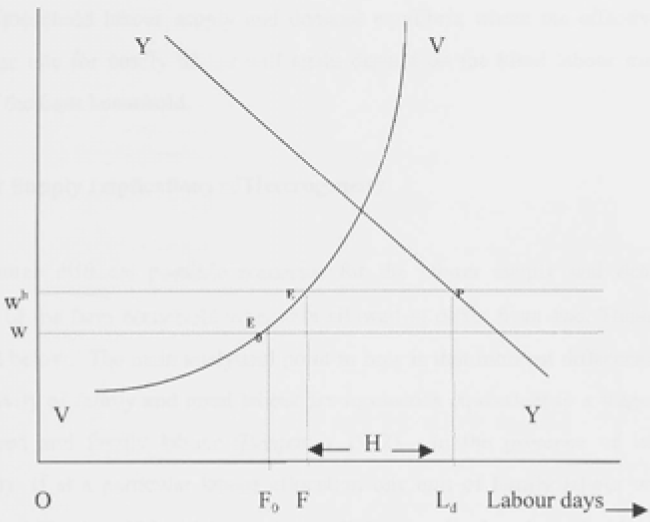
[16] The separability of the labour inputs in the production function is not a necessary condition to relate the shadow wage rate of family labour to the wage rate for hired labour. However without the separability property the ratio of the ratio of the marginal products of family and hired labour are functions of the level of other non-labour inputs as well; so this relationship will be less tractable.

Under non-separability of the production function given by Eq. 3.3, $\theta^* = h'(F^*, G^*, A^*)$ where A is the land input. Although the shadow wage rate for family labour (w*) can be related to w^h and θ^*, the value of w* will vary among farmers who hire in labour at the same wage rate but whose farm size differ. When the production function is separable, θ^* will be independent of A and other non-labour inputs. Hence w* can be related directly to w^h and the levels of F and H only.

The farm household labour demand and labour supply equilibrium with $w^h > w$ (implying $M = 0$) and $\theta^* = 1$ is illustrated in Figure 3.3 below. As in Fig. 3.1 and 3.2, the VV curve plots the rising real cost of family labour which is the utility cost of leisure foregone represented by U_{ℓ}/U_C. The YY curve is the marginal product of aggregate labour drawn under the assumption that family and hired labour are homogeneous inputs ($\theta = 1$, everywhere). Since at the margin hired labour is used, the aggregate labour demand equilibrium is represented by the point P where the marginal product of the last unit of aggregate labour equals the hired wage rate. L_d measures the total labour days utilized on the farm which under the assumption that $\theta = 1$ everywhere is simply the sum of F and H in natural units. In the presence of hired labour which is perfectly substitutable with family labour, the family labour supply equilibrium is represented by point E where U_{ℓ}/U_C equals w^h, and not the point E_0 where U_{ℓ}/U_C equals w.

When $w < w^h$ and $\theta = 1$, the market wage (w) that family labour could earn when working off farm is *irrelevant* to the labour supply equilibrium of the big farm household that is hiring in labour at the higher wage rate, w^h. For every extra unit of family labour applied between E_0 and E, the real cost of labour is below w^h while the gain in consumption is w^h - the value of the wage payment for hired labour that is foregone when family labour displaces hired labour on a one to one basis. Hence the farm household attains a higher level of utility at Point E that at point E_0 (as long as the point E is feasible, with total work- days less than or equal to the total time endowment, T).

Figure 3.3: Labour Supply and Demand Equilibrium for a Big Farm Household



In the case of a small farm household the relevant family labour supply choice is about the allocation of total time worked into the own farm labour input (F) component and the off-farm market employment (M) component at wage w. Here also, all the relevant scenarios can be derived from Eq. 3.23. Again, when $\theta = 1$ everywhere, if $w > w^h$ one would observe complete specialization in the labour allocation of the small farm household also. All of its labour supply will be on market employment at wage w, and all of the labour input on the family farm will be hired labour paid a wage w^h per unit. Eq. 3.21 is satisfied with $\mu_3 = 0$ and μ_1 and μ_2 non-negative. If $w < w^h$ then hired labour will not be used and family labour is allocated to the family farm till the marginal product of family labour is equated to the market wage rate. The remaining work days are devoted to market wage employment. In either case the wage rate which applies to the determination of the household labour supply equilibrium is the market wage rate. The optimality condition derived from Eq. 3.23 in this situation is exactly equivalent to the equilibrium conditions given by Eq.3.15 derived in Section 3.2 where family and hired labour are treated as completely distinct inputs.

The more general case when θ is different from one give rises to a wide variety of alternative household labour supply and demand equilibria where the effective or shadow wage rate for family labour will again depend on the hired labour market exposure of the farm household.

3.5 Labour Supply Implications of Heterogeneity

There are three different possible scenarios for the labour supply and demand equilibrium of the farm household when θ is allowed to differ from one. These are summarized below. The main analytical point to note is that inherent differences in the productivity of family and hired labour are in essence equivalent to a wage gap between hired and family labour (Benjamin 1992). In the presence of labour heterogeneity, if at a particular labour allocation one unit of family labour would substitute for $1/\theta$ units of hired labour, the implicit price the farm household would place on one unit of family labour would be w^h/θ, if hired labour was also being used by the household.

Case(i) : if $w > w^h/\theta^*$

All farm households are completely specialized in their labour allocation. All days of family labour supply (L_s) are devoted to wage employment in the local labour market (at wage w); and all labour input (L_d) on the family farm consists of hired labour only (paid a wage w^h).

In this scenario it is not meaningful to distinguish small farm households (who are net sellers of labour) or big farm households (who are net buyers of labour). The optimal labour supply and demand conditions for all households, irrespective of their total labour demand needs, are identical and are given by

$$(3.25) \qquad U_{\ell}/U_C = w \qquad L_s = M\ ;\ (F=0)$$

$$(3.26) \qquad p\frac{\partial Q}{\partial H} = w^h \qquad L_d = H$$

These optimal conditions are independent of whether θ is less than, equal to or greater than one. When $w > w^h/\theta$, irrespective of the value of θ, it pays the farm household to specialize by offering all of its desired labour supply on the off-farm labour market, while hiring in all its labour demand requirements.

In this scenario, in spite of the presence of labour heterogeneity ($\theta \neq 1$), the consumption/labour supply component of the farm household's decisions are independent from the production side choices for all households. The effective wage rate which determines the farm household's optimal labour/leisure choice is the market wage rate, w; and the farm household model is recursive.

Case(ii) : if $w = w^h/\theta^*$

When the differences in the wage rates for family and hired labour exactly offset the difference in marginal productivity, the equilibrium conditions are given by:

$$(3.27) \qquad w^h = p\frac{\partial Q}{\partial H}$$

$$(3.28) \qquad U_{\ell}/U_C = w = \frac{w^h}{\theta^*} = p\frac{\partial Q}{\partial F}$$

When any observed efficiency difference between family and hired labour is exactly mirrored in the difference in the market wage rates available to hired and family labour, the family and hired labour inputs become perfect substitutes (in the neighborhood where $w = w^h/\theta*$). This sub-case reduces to the conventional farm household model where family and hired labour are treated as homogenous inputs with a common market wage rate. The only adjustment is in the units of measurement - i.e. whether total labour is measured in units of family days or hired labour days with $\theta*$ being the conversion factor for the quantities and the wage rates.

In this scenario, as in the case of homogeneous labour inputs, the farm household model cannot be solved for specific levels of F, H or M. What is determined is total labour demand in effective units ($F + \theta*H$); but its composition between F and H is indeterminate. Similarly on the consumption/leisure component of the model, only total labour supply days ($F + M$) is determined; its allocation between F and M is arbitrary. In this situation altering the mix between F and H and between F and M, while keeping total labour supply and demand fixed has no effect on farm profits or wage income, and hence no effect on the household utility level.

Case(iii) : if $w < w^h/\theta*$

In this scenario the analytical structure of the farm household differs according to the net position of the household in the hired labour market. Three distinct farm household types can be distinguished. For each type the effective wage rate that determines the labour supply equilibrium will differ.

(a) the small farm household whose total labour supply will consist of work on the family farm (F) supplemented by off-farm market wage employment (M) at wage w. The effective wage rate for this household is the market wage rate, w, and its equilibrium position is characterized by

$$(3.29) \qquad U_{\mathcal{L}}/U_C = w = p\frac{\partial Q}{\partial F} \qquad M > 0;\ F > 0;\ H = 0$$

These conditions are exactly equivalent to the equilibrium conditions for a household with M > 0 given by Eq.3.15 in Section 3.2 where family and hired labour are treated as completely distinct inputs. From Eq. 3.29, this sub-model is also recursive.

No hired labour is used because the total family labour supply exceeds the total labour demand and the wage cost of family labour is lower than that of an equivalent unit of hired labour. It will be more profitable for the farm household to use family labour valued at wage w than hired labour at an effective wage of w^h/θ.

(b) <u>the autarchic household</u> which neither sells any family labour on the market nor employs any hired labour (i.e. M = H = 0). Its equilibrium is described by

$$(3.30) \qquad \frac{U_\ell}{U_C} = w^* = (p\frac{\delta Q}{\delta F})$$

$$(3.31) \qquad \text{where} \quad w \leq w^* \leq \frac{w^h}{\theta^*}$$

The effective wage rate, w*, faced by this household is indeterminate within the bounds specified above.[17] Since w* is partly a function of the marginal product of family labour which depends on production technology, the model for the autarchic farm household is not recursive.

---

[17] These bounds on w* are slightly wider when $\theta* > 1$. If hired labour has a higher marginal product the farm household can be autarchic only if there are no net gains from transferring one unit of family labour from own farm work to market work, while replacing the unit of family labour with hired labour for farm cultivation. This leaves unchanged the total labour supply of the farm-household, hence the net gains can be determined by the change in consumption which is given by

$$(N3.17a) \qquad (w - p\frac{\partial Q}{\partial F}) + (p\frac{\partial Q}{\partial H} - w^h)$$

A farm household can remain autarchic only if the above expression is non-positive. Noting that $w^* = p\frac{\partial Q}{\partial F}$ and $p\frac{\partial Q}{\partial H} = \theta^* p\frac{\partial Q}{\partial F}$, the required condition necessary for an autarchic state to be optimal reduces to

$$(N3.17b) \qquad w \leq w^* \leq \frac{w^h - w}{\theta^* - 1} \quad \text{when } \theta* > 1.$$

It can be readily shown that $\frac{w^h - w}{\theta^* - 1} > \frac{w^h}{\theta^*}$ if, as assumed, $\theta* > 1$ and $w < w^h/\theta$.

(c) the big farm household which must supplement its family labour days with hired labour in order to meet the total demand for labour in its farm cultivation. Given that $w < w^h/\theta$, the big farm household never supplies any labour to the market at wage w since that same unit of labour when applied to the family farm earns a return of w^h/θ when family labour can substitute for $1/\theta$ units of hired labour.

The optimum labour input allocations will satisfy

$$(3.32) \qquad w^h = p\frac{\partial Q}{\partial H}$$

$$(3.33) \qquad U_{\ell}/U_C = w^* = p\frac{\partial Q}{\partial F}; \qquad F > 0;\ M = 0$$

$$\text{where } w \leq w^* = \frac{w^h}{\theta^*} \quad \text{since by assumption}$$

$$(3.34) \qquad \frac{w^h}{w^*} = \left(\frac{\partial Q/\partial H}{\partial Q/\partial F}\right)_{F^*H^*} = \theta^*$$

Consequently, for a household which uses hired in labour the shadow wage rate w* which determines its own labour supply equilibrium is not "subjective" any more as it was in the model with family and hired labour as distinct inputs.[18] With separability this wage rate can be related directly to the wage rate for hired labour with the adjustment for the difference in the marginal productivity, if any, of the two types of labour at that particular equilibrium. Hence, $w^* = w^h/\theta^*$.

---

[18] From Eq. 3.16 $w^* = w + \frac{\mu_1}{\lambda}$ when family labour is modeled as distinct input and all family labour is supplied on the family farm. Since μ_1 and λ are Lagrange multipliers that vary with the constraints of the model, the effective wage rate w* is subjective and cannot be identified from the observed market wage rate w.

The farm household sub-model for the big farm household that hires in labour is not generally recursive since the effective wage rate for family labour, w^h/θ, is not exogenously given to the household. It will depend on the allocation of F and H which affects θ^*. However, if within the relevant range of labour allocations for a particular household, θ^* happens to be constant then the sub-model for the big farm household is recursive. In this case the household equilibrium will be identical to the case of a household facing an exogenously given market wage rate of w^h for hired labour and a market wage rate w^h/θ^* for family labour.

Diagrammatic illustration for Case (iii) when $w < w^h/\theta^*$

The labour supply and demand equilibrium for these three types of households are illustrated in Figure 3.4 for a specific scenario where $\theta^* < 1$ and $w < w^h < w^h/\theta$. The marginal product of family labour is denoted by the Y_F schedule while the marginal product of hired labour is denoted by the Y_H schedule. As drawn Y_H is always below Y_F indicating that at any given labour allocation of F and H, adding one more unit of F leads to a higher increase in output than adding an extra unit of H. This implies $\theta < 1$ in the entire range of total labour demand.

In each panel the point labeled **E** denotes the appropriate labour supply equilibrium. Hence, in estimating the labour supply component of the farm household it is important that the wage rate variable to be used in the regression analyses conforms to the wage rate relevant for point E. This wage rate varies for the three different household types[19] as indicated in Panels (a), (b) and (c) of Figure 3.4. The effective wage rate for family labour is higher than the market wage rate in the autarchic and big farm households.

[19] Fig 3.4 is drawn with $w < w^h$. This is not a necessary condition for the three different household types to be defined. The necessary condition is $w < w^h/\theta$. When $\theta < 1$, this condition can be satisfied even if $w > w^h$.

Panel (a) shows that the differences in the productivity of family and hired labour have no effect in neighborhood of the labour supply equilibrium of the small farm household. At the margin it faces a given wage rate which determines the total amount of its labour supply. The Point E in Panel (a) is unaffected by the value of θ. The only relevance of productivity differences in Panel (a) is that with $\theta < 1$ and $w < w^h/\theta$, the small farm household does not simultaneously use hired labour as well as supply its own labour in market work.

In Panel (b) the farm household is autarchic: it neither buys any hired labour nor does it sell any of its family labour on the market. Total household labour supply exactly matches the total demand for labour in farm production (the amount OF). The household acts as if it were facing an effective wage rate w^* which equates the demand and supply of family labour. This household does not sell any of its family labour on the market because the returns to labour on own farm cultivation are higher than the market wage rate. On the other hand, given the family labour input of OF, the marginal product of an extra unit of hired labour is less than the hired labour wage rate w^h.

As drawn in Panel (b) the autarchic equilibrium is created by the fact that $\theta < 1$. The wage gap between w and w^h by itself is not sufficient to create an optimal autarchic equilibrium since $w^* > w^h$ as drawn. If $\theta = 1$ and the marginal productivity of both family and hired labour were to be represented by the Y_F schedule, it would not be optimal for this household to be autarchic.[20] In that case family labour supply to own farm cultivation would only be OF^0 (instead of OF) and a few extra units of labour would be hired in (upto the point of intersection of the Y_F schedule and the w^h wage line).

Finally in Panel (c), the effective wage rate applied to family labour at point E is directly affected by the value of θ. The fact that $\theta < 1$ has two separate effects on the optimal labour supply and demand choices of the big farm household:

[20] With $\theta = 1$ an autarchic equilibrium could still be optimal if the intersection of the Y_F and V schedules was such that $w < w^* < w^h$.

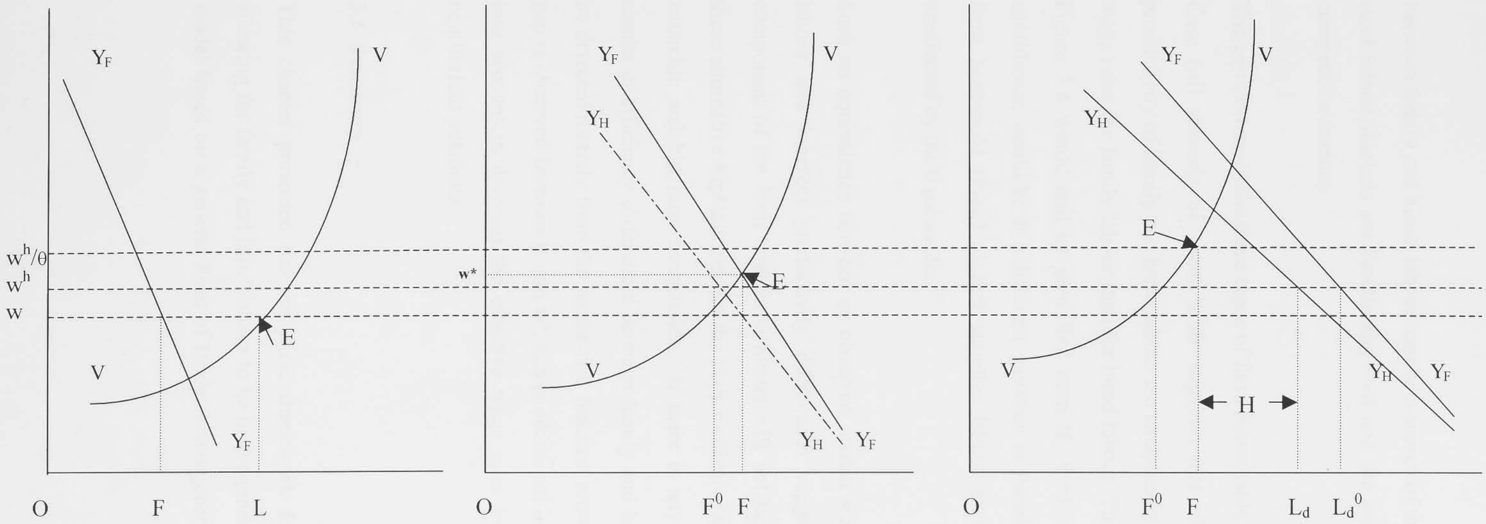
(i) the first effect is that total labour demand is reduced (from OL_d^0 to OL_d). Without any differences in the productivity of the two types of labour, the Y_H schedule would coincide with the Y_F schedule and the optimal total labour demand would be the point of intersection of the labour marginal productivity schedule with the w^h wage line (which gives total labour demand of OL_d^0 .

(ii) the second effect is a higher labour supply response for family labour because it leads to a higher effective wage rate for family labour. Family labour supply is OF with $\theta < 1$ as compared to OF^0 with $\theta = 1$.

The combined effect of a lower total labour demand and a higher proportion of it coming from family sources means that the demand for hired labour is considerably reduced on big farms because of the efficiency differences. The amount of hired labour used on the big farm household illustrated in Panel (c) is the distance between F and L_d (indicated as H). If family and hired labour were homogeneous inputs the demand for hired labour on this big farm would be equal to the distance between F^0 and L_d^0 instead.

As discussed in Chapter II, one of the key "stylized facts" about agricultural production in developing countries is that small farms are cultivated more intensively with higher levels of variable inputs used per hectare than bigger farms. This often leads to the celebrated observation of an inverse relationship between farm size and output per hectare. The relatively greater application of inputs on smaller farms is most pronounced in the case of labour. Per hectare labour input on small farms is consistently higher than on bigger farms over a large range of farm sizes; and this result holds whether or not average yields on small farms are higher (Berry and Cline, 1979). A traditional explanation for the higher labour intensity on small farms has focussed on factor market imperfections (Sen 1966, 1975). By comparing Panels (a) and (c) of Figure 3.4, it is clear that efficiency differences

Figure 3.4: Farm Household Equilibrium with Labour Heterogeneity



Panel a : Small Farmer	**Panel b**: Autarchic	**Panel c**: Big Farmer

Equilibrium Conditions :

Panel a:

$$\frac{U_{\ell}}{U_C} = w = (p\frac{\partial Q}{\partial F})$$

$H = 0$; $M > 0$

Panel b:

$$\frac{U_{\ell}}{U_C} = w^* = (p\frac{\partial Q}{\partial F})$$

$w < w^* < w^h/\theta$

$H = M = 0;$

Panel c:

$$\frac{U_{\ell}}{U_C} = w^h/\theta = (p\frac{\partial Q}{\partial F})$$

$M = 0;$

$$(p\frac{\partial Q}{\partial H}) = w^h$$

between family and hired labour can be a source of the inverse relationship even if rural labour markets are functioning well and the market wage rates are set in a competitive manner.

The equilibrium of the three types of farm households illustrated in Fig. 3.4 for the Case (iii) scenario of $w < w^h/\theta$ suggests that efficiency differences in the productivity of family and hired labour are analytically equivalent to a lower market wage rates for family labour than for hired labour. The three types of household in Figure 3.4 would still be identified even if $\theta = 1$ as long as $w < w^h$, but the equilibrium would be at a different position in Panels (b) and (c). Only the small farm household (Panel a) that supplies labour to the market wage rate w is unaffected by the θ parameter.[21]

Such an equivalence between an observed market wage gap for family and hired labour and inherent productivity differences suggests that the labour supply component of the farm household model will not be able to distinguish between these alternative explanations for the higher effective wage rate for family labour in autarchic and big farm households. If there is any labour heterogeneity which results in efficiency differences between family and hired labour ($\theta^* \neq 1$), it must be detected directly from the production function estimations. If in addition a wage gap is observed between w and w^h, this is additional information that must be taken into account in deriving the effective wage rates to describe the labour supply equilibrium positions.

3.6 Summary

This chapter presented the analytical framework for a farm household model allowing for family and hired labour to be heterogeneous inputs. A farm household model based on a general form of labour heterogeneity that allows for family and hired labour (as well as other relevant labour categories) to be completely distinct inputs is analytically unappealing and empirically difficult to estimate. A particular problem in treating family labour as a distinct input is that the wage rate for determining the labour/leisure equilibrium choice of the household is unidentified when family labour is devoted solely to own farm cultivation.

An alternative approach to modeling labour heterogeneity issue with a composite effective labour input is more appealing. With such a specification, the effective wage rates for family labour can be related directly to the observed market wage rates at which non-family labour can be hired in. The aggregation of various labour inputs, however, requires that the labour inputs be separable from the other inputs in the production function. If this condition is met the ratio of the marginal products of family and hired labour become independent of the levels of the other non-labour inputs. Consequently, the first order conditions for the optimal use of hired labour can be used to derive the effective wage rates for family labour as well. This framework is general enough to accommodate wage gaps that arise when the hiring in wage rate and hiring out wage rates for family labour differ.

The separability of the labour inputs in the production function in essence makes family labour equivalent to a traded input even when family labour is supplied only to own farm production work. Such an equivalence is obtained by comparing the marginal product of family labour with that of the traded input (hired labour). This relationship can be used to derive the effective wage rate for family labour on large farms which also employ hired labour. This considerably simplifies the estimation of the labour supply component of the farm household model for households engaged in own farm work only.

The predictions made by a model of the farm household with heterogeneous labour inputs are testable. When $w > w^h/\theta$, it should lead to complete specialization in the

[21] This ignores potential general equilibrium effects whereby the demand for hired labour is affected by the θ parameter, and hence has an effect on the equilibrium market wage rate at which a small farm household can work off-farm.

labour supply and demand of all farm households. This situation, however, is rarely observed in practice. In the more likely scenario of $w < w^h/\theta$, it gives rise to three distinct types of farm households: the small farm household that does not employ any hired labour, but sells surplus family labour on the market, an autarchic household which neither buys nor sells any labour, and a big farm household where all of family labour is devoted to own farm cultivation, supplemented by hired labour. The effective wage rate which identifies the labour/leisure equilibrium for all three household types can be identified, and an estimation strategy devised based on observed market wage rates and estimated parameters of the production function.

CHAPTER IV

ESTIMATION STRATEGY

This Chapter develops the econometric strategy for both detecting and incorporating potential labour heterogeneity in the production and labour supply component of the farm household model. The model estimation results, based on household data from the *tarai* region of Nepal, are presented in Chapters VI and VII. In this Chapter Section 4.1 reviews the general estimation issues involved and briefly discusses alternative strategies. Section 4.2 outlines the sequential estimation strategy proposed by Jacoby (1993) for estimating non-recursive farm household models. Section 4.3 presents an adaptation of this methodology to the estimation of a farm household model with labour heterogeneity that is consistent with the structure developed in Sections 3.4 and 3.5 of Chapter III. Section 4.4 contains a brief discussion of the required adjustments to the variance-covariance matrix in a sequential estimation strategy.

4.1 Estimation Issues

The conventional approach to the empirical estimation of farm household models treats family and hired labour as homogenous inputs valued at a common wage rate. As a consequence a two step estimation procedure is feasible. In the first step, the production side of the model is estimated through either a production function (Barnum and Squire 1979), a profit function (Lau, Lin and Yotopoulos 1987), or a cost function (Binswagner 1974). These are specified for a given level of commodity dis-aggregation for outputs and inputs. The labour input variable is in terms of total labour: its composition between family and hired labour need not be addressed. This first step results in a set of output supply and input demand equations that are functions of input and output prices, and of farm characteristics, including fixed inputs.

In the second step, a consumer demand system is estimated to represent the farm household's preferences. This second component can be modeled either as a complete dis-aggregated demand system (such as the Linear Expenditure or Log Linear Expenditure system) that specifies leisure as one of the consumption goods (Kuroda and Yotopoulos 1980). An alternative approach is to specify directly a reduced form labour supply equation (Rosenzweig 1980) which can be made consistent with a demand system with leisure and an aggregate consumption good.

Irrespective of the approach and specific functional forms chosen for production and consumption components of the model, the distinctive feature of the conventional estimation strategy is that the consumption/labour supply choices are estimated independently of the production side of the model, relying on the recursive property of the standard farm household model. The only connection between the production and consumption side in the conventional estimation of the recursive model is that the consumption choices are conditioned on the level of "full income" of the household, which includes the value of farm profits at the production optimum.[1] On the consumption side, the commodity demands and labour supply are functions solely of commodity prices, including labour wage rates, household full income, and possibly some household and individual characteristics which affect preferences. Hence the full consumer demand system or the labour supply equation alone can be estimated in an independent fashion based solely on exogenous variables and the farm profit variable which can be computed from the observed production side choices (Singh, Squire, and Strauss 1986c1: 20).

On the other hand, when family and hired labour must be treated as heterogeneous inputs, in the general case the recursive structure of production and consumption choices breaks down.[2] Econometric estimation of fully non-separable models is analytically cumbersome as well as very time-intensive to implement. Joint

---

[1] The full income of the farm household includes the imputed short run farm profits which represents the pure returns to the land input. (See Equation 3.2n in Section 3.2 of Chapter III). The method of imputing the value of farm profits already reflects the labour supply choices of the farm household, and hence the consumer demand side estimation is conditional on this specific level of full income.

[2] The recursive structure of the farm household model may break down for other reasons as well - such as if home produced consumer goods are imperfect substitutes for market purchased consumer goods or if the household assigns different levels of dis-utility to work days applied to own farm cultivation and work days on the off-farm labour market. Lopez (1984) is an example of a non-recursive model based on varying preferences between work on the family farm and off-farm work.

estimation of farm production, input demands as well as consumer demand equations may not even be feasible at times. It may be impossible to solve for the reduced forms of these equations analytically (Singh, Squire and Strauss 1986c1: 22). Given that a non-separable farm household model would be highly non-linear in the parameters it may not be feasible to identify the structural parameters of interest even when it is possible to solve for their reduced forms (Jacoby 1993). A joint-systems estimation relying on full information methods will be very cumbersome, especially as the level of commodity dis-aggregation increases. Moreover the data requirements for the joint estimation approach are enormous. Data is required for all endogenous variables in the system (in addition to the set of properly exogenous variables) while in a recursive model estimation can be done with limited data sets which cover only specific components of the model.

While non-recursive models are difficult to implement empirically, there are costs to assuming a recursive estimation strategy when the underlying model is genuinely non-recursive. The inter-dependence of the production and consumption choices of farm household models affects empirical estimation based on an assumed recursive structure in two ways.[3] The usual parameters estimated in output supply, factor demand and consumer demand systems that incorrectly assume a recursive structure are statistically inconsistent. Secondly, these parameters by themselves are unable to identify the full effect of the comparative statics of the truly non-recursive model because the latter model will have additional terms that have not been estimated.[4] The total error resulting from the combination of the inconsistent estimates and the missing terms in the comparative static derivations will be difficult to gauge.

---

[3] The subsequent discussion in this paragraph closely follows Singh, Squire and Strauss (1986c2: 48).

[4] For instance, in deriving the comparative statics of a change in the price of farm output on household leisure demand, in the recursive model with a fixed wage rate for family labour the output price change has only an income effect because it changes the level of farm profits. But in a non-recursive model with a subjectively determined shadow wage rate for family labour, the change in the output price has an additional effect on leisure demand. The household -specific shadow price for family labour that equates family labour supply and demand will also change in response to the change in the output price. The latter term - the marginal effect on the shadow wage for family labour due to a change in the output price will not have been estimated in a recursive model structure. See Strauss (1986) in Singh, Squire and Strauss (1986) for the comparative statics of a farm household model involving shadow or virtual prices.

Previous studies have not adequately looked into the question of how large the errors are in implementing a recursive estimation strategy when the true underlying model is non-recursive. Lopez (1984) - which is also summarized in Lopez (1986) - is the only example where estimation results of a full information method based joint estimation strategy have been compared with the results from assuming the conventional recursive farm household model structure. His results show significant difference in the values of the parameters of the farm household model when using a recursive *versus* a non- recursive estimation strategy. For instance, the estimates of the own wage elasticity of family labour supply changes from 0.19 in a recursive estimation strategy to 0.04 in a non-recursive strategy. Unfortunately, Lopez does not report standard errors for each set of estimates so one cannot test whether the differences in the estimated value are statistically significant or not. A second and more important issue which was not addressed is whether the differences in these two estimates are also of economic significance -i.e., are they different enough to lead to very different results on specific applications of the model, especially on implied policy prescriptions. Again, there is little empirical evidence on how substantially significant such model errors can be.

4.2 The Two Step Estimation Strategy for a Non-Recursive Model

The possibility that such errors could be significant - statistically and economically - but not always large enough to warrant a full-scale joint estimation has led to the development of alternative estimation strategies. The objective has been to develop estimation procedures that are valid for a non-recursive model specification but which retain the tractability of the step-wise estimation of the production and consumption/labour supply components of the conventional farm household model.

One such approach has been proposed and applied by Jacoby (1993). Jacoby's approach is a general methodology for estimating a structural labour supply equation for workers who are self-employed. It is analogous to the treatment of labour supply in the presence of progressive income taxes that was pioneered by Hall (1973) and which involves "linearizing" an underlying non-linear budget constraint. This approach is widely used in the labour supply estimation in developed country settings,[5] but Jacoby's application to a farm-household setting is novel.

In the context of farm-households whose members are purely self-employed ($M = 0$ in the terminology of Chapter III), the household budget constraint for consumption is given by the net farm output which is an increasing function of family labour days applied to farm production. Consequently, the household faces a concave budget constraint in the form of a well behaved agricultural production function with diminishing marginal product of family labour. The methodology proposed by Jacoby is to "linearize" the budget constraint at the household's optimum point. The objective is to convert the farm-household labour/leisure choice with a concave budget constraint into an equivalent standard consumer problem where the farm household acts as if it were facing a linear budget constraint with a fixed wage rate for its labour but which leads to the same optimal labour supply position. This "linearization process" is illustrated Figure 4. 1.

For the simple farm household model presented in Chapter III where labour and land are the only production inputs, from Eq. 3.2, the household's budget constraint for consumption is given by

$$(4.1) \qquad C = (pQ - w^h H) + wM + E$$

The three components of the budget constraint represented above are the net returns from own farm cultivation ($pQ - w^h H$), wage income when family labour works on the off-farm labour market (wM), and non-labour endowment income (E) which is exogenously given. The full range of such a budget constraint for the farm household is diagrammatically represented by the solid bold line OBPQ in Figure 4.1.

The OBPQ line has three segments to match the three components of the household's budget constraint in Eq. 4.1. The vertical segment OB measures the

---

[5] Burtless and Hausman (1978) and Blomquist (1983) are early examples.

level of consumption that can be funded solely from the household's non-labour endowment income. The distance OB measures the value of the variable E in Equation. 4.1 above. The curved segment BGP represents net income from farm production.[6] It includes the returns to the ownership of the land as well as to family labour applied to own farm production. As drawn, BGP is a part of the farm production function for increasing levels of family labour application (holding all other inputs constant) which is represented in its entirety by BGQ. Hence, the slope of BGP is the marginal product of family labour which is decreasing as more family labour is allocated to farm production. The final linear segment PK represents labour income from off-farm market work. The slope of PK is the exogenously given market wage rate, w.

Given the household budget constraint line OBPQ and household preferences over leisure and consumption denoted by the indifference curves I_1 and I_2 (and assuming that at least some family labour is devoted to on-farm production),[7] two equilibrium positions are possible. The tangency of the indifference curves can either be with the curved section BP of the budget constraint or on a linear section such as PK.

The scenario that is drawn in Figure 4.1 is of a farm household equilibrium on the curved section BP of the budget constraint. Given household preferences, the production technology and the market wage rate (w), the household finds it optimal to supply all its labour to own farm production and not engage in any off-farm work. This optimal labour allocation is denoted by the point G. At this equilibrium the subjective valuation of family labour (U_{ℓ}/U_C) is set equal to the marginal product of family labour. The latter is given by w*, the slope of the farm production function at G. Hence the optimal allocation of family labour denoted by the point G is

---

[6] Since the price of the consumption good $\mathbf{p}_c$ was normalized to one, the value of farm output can be incorporated directly into the household's consumption budget constraint without distinguishing real and nominal consumption quantities.

[7] This avoids corner solutions involving complete non-participation of family labour in productive activity. The labour force participation decision can be modeled in alternative ways, for instance, with a probit model.

supported by a "shadow wage rate", w*, that is necessarily greater than or equal to the market wage rate, w.[8]

The process of "linearizing" the household budget constraint at the optimal point G is also illustrated in Figure 4.1. If the concave budget constraint OBPQ that the farm household actually faced would have been replaced by an artificial linear budget constraint represented by the dashed bold line DGN in Figure 4.1, the farm household equilibrium remains unchanged at point G. This artificial linear budget constraint DGN is uniquely defined by two parameters: (i) the slope w*; and (ii) the intercept OD. The intercept OD represents the consumption level that could be attained, based on the fictitious linear budget constraint line DN, if all available family labour time was devoted to leisure.

Hence, if the budget constraint of this particular farm household was re-formulated in a linear form, uniquely defined by the slope w* and intercept OD, the equilibrium point G would represent the optimum solution of the standard consumption-leisure choice problem faced by a household that was not involved in its own farm production. G represents the labour supply equilibrium of a household with a fixed labour endowment T and an exogenously given non-labour endowment income (equal to the amount OD), where the household faces an exogenously given market wage rate of w*, which it takes as given for all levels of its labour supply.

That is, the point G is also the solution to the problem

(4.2) $\max \; U(C, \mathcal{L})$

subject to

(4.3) $C = w^*(T - \mathcal{L}) + E^*$

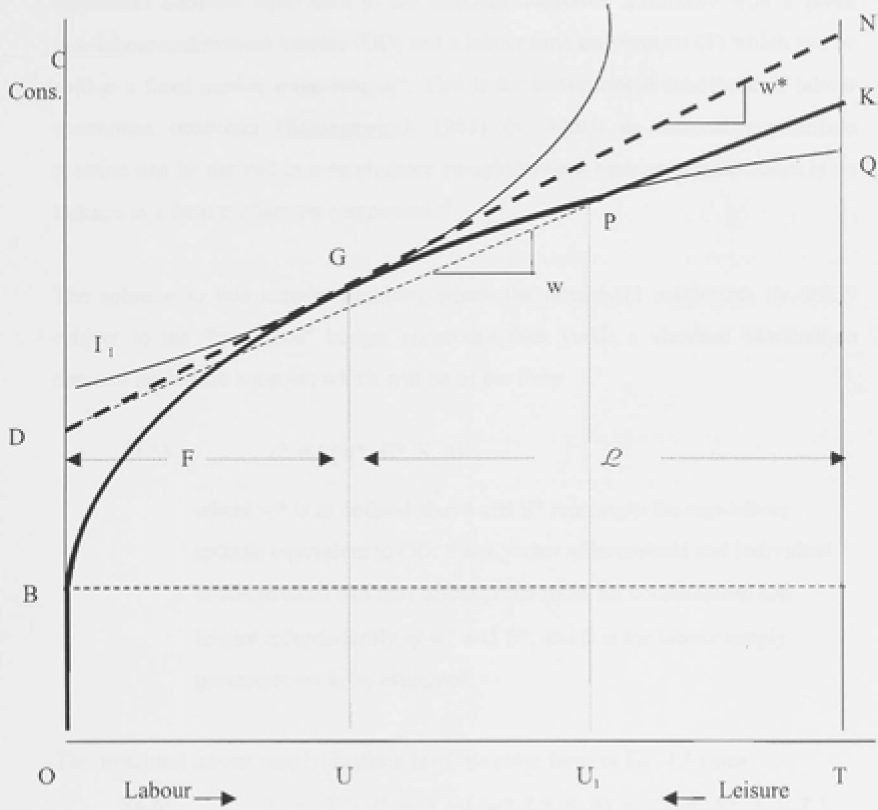
for which the first order condition

(4.4a) $U_{\mathcal{L}}/U_C = w^*$

---

[8] Figure 4.1 is identical to Panel A in Figure 3.1 of Chapter III, except for the extension of the vertical axis to include the BO segment of the budget constraint. Point G in Figure 4.1 is identical with point G in Figure 3.1(A).

Figure 4.1: Linearized Budget Constraint for a Farm Household Equilibrium



is identical to the first order condition for the labour supply component of the farm household's utility maximization problem (the first equality in Eq. 3.16) where[9]

(4.4b) $$w^* = w + \frac{\mu_1}{\lambda}$$

---

[9] λ is the Lagrangian multiplier associated with the farm-household utility maximization problem given in Chapter III by Eq. 3.6, and μ_1 is the complementary slackness parameter associated with the non-negativity of off-farm labour supply (the M variable) as specified in Eq. 3.7a.

This re-formulation of the consumer problem with a linearized budget constraint in essence *takes the farm out of farm-household's decision-making locus*. The farm household becomes equivalent to the standard consumer household with a given non-labour endowment income (OD) and a labour time endowment (T) which can be sold at a fixed market wage rate, w*. This is the conventional household of labour economics textbooks (Killingsworth 1983) for which an internal equilibrium solution can be derived in a much more straightforward manner because there is no linkage to a farm production component.[10]

The solution to this simpler problem, where the household maximizes its utility subject to the "linearized" budget constraint then yields a standard Marshallian demand for leisure equation which will be of the form

(4.5) $\mathcal{L} = \mathfrak{f}(w^*, E^*, S, \beta)$

where w* is as defined above and E* represents the non-labour income equivalent to OD. S is a vector of household and individual characteristics that may affect preferences for consumption and leisure independently of w* and E*; and β is the labour supply parameter set to be estimated.

The structural labour supply function is of the same form as Eq. 4.5 since

(4.6) $L_s = T - \mathcal{L} = T - \mathfrak{f}(w^*, E^*, S, \beta) = \ell(w^*, E^*, S, \beta, T)$

where L_s is the total labour supply of the household,[11] and T is conventionally treated as an exogenously given constant.

[10] The only difference between the textbook version of a labour household equilibrium and the farm household equilibrium with a linearized budget constraint is in the interpretation of the full income of the household. In the former the full income is completely determined by E and w*T and does not reflect any household choices. In the latter the full income of the farm household is conditioned on the observed demand for leisure, and conversely the optimal leisure/labour choices are also conditioned on the production choices determining the level of farm profits (Jacoby 1993:906).

[11] L_s includes all components of the household labour supply including other productive activities such as home processing and cottage industry activities. At the optimum, the marginal returns to all productive activities should be set equal to each other. It is not required that the shadow wage rate be defined only in terms of the marginal product of family labour in farm production. The marginal returns to other household production processes could be used as well, although it would normally be more difficult to estimate the production function for these other activities.

The two variables (w^* and E^*), needed to identify the linearized budget constraint and hence to allow the estimation of Eq. 4.6, can be derived directly from a prior estimate of a farm production function. As noted before, w^* is just the marginal product of family labour in farm production at the optimum production point. E^* is the level of an assumed endowment income which includes the level of farm profits at the production optimum. This level of farm profits is calculated by deducting all labour cost and other variable costs from the gross value of farm production. The labour costs for family labour are based on the assumption that family labour is paid a fixed wage rate of w^* for each unit of family labour applied to the farm. That is,

(4.7) $$E^* = \pi^*(w^*) + E$$

where π^* = "shadow" farm profits -the maximized value of profits at the equilibrium point G using a shadow wage rate, w^*

$$= \max_{F,H} \{ p\,Q(F, H, A) - w^h H - w^* F\}$$

(4.8a) $$= \{ p\,Q(F^*, H^*, A) - w^h H^*\} - w^* F^*$$

(4.8b) $$= Y^*_N - w^* F^*$$

(Y^*_N represents the maximized total returns to farm cultivation net of all purchased inputs, but including the implicit return to family labour applied to own farm work; and F^* and H^* are the optimum levels of family and hired labour input).

Since E^* depends only on variables that are either exogenous or depend only on w^*, the only variable needed to linearize the farm household's budget constraint is w^*.

In summary, the application of the "shadow wage" method for estimating a labour supply equation for farm households engaged only in self-employment consists of three steps:

(i) estimate a whole-farm agricultural production function in which family labour is recognized as a distinct input

(ii) using the estimated parameters of the production function, derive the marginal product for family labour for each sample household. Using this estimated value of the marginal product as the effective wage rate (w^*) for family labour, derive the shadow farm profits at the optimum production point. Then compute the total non-labour endowment income variable E^* as the sum of shadow farm profits and the normal non-labour endowment income E.

(iii) estimate a standard labour supply function of the form of Equation 4.6 by regressing total household labour supply on w^*, E^* and S, assuming the household treated these shadow wages and shadow endowment income as being exogenously given to the household *for all levels* of its labour supply.[12]

The labour supply equation in (iii) can be estimated independently or as part of a joint demand system with the other consumption goods of the model.[13]

This methodology of linearizing the concave budget constraint of the farm household works just as well for households participating in the off-farm labour market. When $M > 0$ the consumption equilibrium position occurs along a linear section of the household budget constraint (such as PK illustrated in Figure 4.1). In this case the household allocates its family labour to own-farm production up to the point where the marginal return from farm production is equal to the off-farm wage

[12] In Figure 4.1 the full range of the non-linear budget constraint OBPK for a farm household which does not supply any labour on the off-farm market can be represented by a single linear budget constraint at the optimum point. This is in contrast to the linearization process in the presence of progressive income taxes which results in several piece-wise linear budget constraints. The analytical method to solve for the equilibrium labour supply position and the econometric procedures for estimating labour supply functions with piece-wise linear budget constraints is considerably more complex (Moffitt 1986) than in the case of the farm household model outlined above.

[13] This depends on whether leisure is assumed to be separable from other goods in the utility function (Deaton and Muelbauer 1980). Alderman and Sahn (1993) is an example of a joint estimation of leisure and commodity demands using an AIDS specification.

rate (w). Since the wage rate is exogenously given, the equilibrium occurs in an already linear section of the budget constraint. It is still necessary, however, to "linearize" the entire budget constraint along the equilibrium point. This is required in order to reformulate the farm-household utility maximization problem as the simpler utility maximization problem of a consumer household with a fixed non-labour income level and a given labour time endowment.

The linearization is done is the same way as in the case illustrated in Figure 4.1. It is necessary to identify the slope and intercept of a linear budget constraint which would give the same consumer equilibrium as the concave budget constraint OBPQ when the equilibrium occurs in a linear section such as PK. The slope of the linearized budget constraint will be the off-farm market wage rate w (the same slope as PK). The intercept, which gives the imputed non-labour endowment income (E*), will be computed just as in Equation 4.7 with w* replaced by w. When M > 0, the value of family labour in deriving farm profits is based on the opportunity cost of family labour on the off-farm labour market. In this case the marginal product of family labour does not have to be separately estimated to carry out the linearization of the budget constraint. All that is required is to derive the intercept term which represents the value of consumption that can be met from non-labour endowment income, including the imputed value of farm profits derived from deducting the cost of family labour valued at the market wage rate for family labour.[14]

[14] The linearized budget constraint when family labour is also supplied on the off farm labour market can also be illustrated with reference to Figure 4.1 itself. In this case the production equilibrium will be given by the tangency of a wage line with slope w with the production function surface BGQ. To derive this point of tangency within Figure 4.1 simply shift upwards the wage line DK till it is tangent to the BGQ line. Let this line of tangency with slope w be represented by D'K' where OD' is the intercept on the consumption axis. Let P' be the point where D'K' is tangent to BGQ. P' must be to the right of the point G as drawn in Figure 4.1 because the tangency line DN at point G has a higher slope by assumption. D'K' represents the linearized budget constraint for the farm household which works on the off-farm labour market. Its consumption equilibrium will be represented by the tangency of the D'K' line with the highest level indifference curve. This point will be further to the left than P' because by assumption the total labour supply of the household exceeds the family labour applied to own farm cultivation when M > 0.

4.3 A Two Step Estimation Strategy with Heterogeneous Labour

The two-step estimation strategy discussed in the previous section is general enough to carry over to the specific framework of a farm household model with heterogeneous labour inputs. We make two adjustments to the specific procedures used by Jacoby (1993). First, we use the farm production function estimation to embed a test for the heterogeneity of family and hired labour, using alternative representations of the nature of the efficiency differences between family and hired labour. Second, the specification of the labour supply equation is not based on the estimated marginal product of family labour but on the equivalent effective wage rates for family labour that can be derived in a theoretically consistent manner using the parameters that describe the extent of labour heterogeneity. As described in Chapter III (Section 3.5) the effective wage rates for family labour are based on the observed market wage rates for family and hired labour and on the ratio of the marginal products of the two types of labour (θ^*), depending on the net labour market position of each household.

We do not model family labour as a separate input distinct from hired labour but use a nested production structure where aggregate labour is a composite function of family and hired labour. The separable labour composite, if supported by the data, still allows for a general form of labour heterogeneity but it considerably simplifies the analytical solution and estimation of the farm household model. In this case the shadow wage rates for family labour in the labour supply equation need not be set to the estimated marginal products of family labour. We can make use of the additional first order conditions that relate the marginal product of family labour to the market wage rate for *hired* labour on large farms that utilize both family and hired labour. This makes it feasible to derive the correct effective wage rates (w or w^h adjusted by θ^*) for family labour depending on the labour market exposure of the particular farm household, as specified in Section 3.5 of Chapter III.

In the first step we estimate the production structure implicit in Equations 3.3(a) of Section 3.4 in Chapter III where family and hired labour form a separable but possibly heterogeneous composite:[15]

(4.9) $$Q = f(\,g\,(F, H, \Theta), V, \alpha) \;+\; u$$

where A is land area cultivated

$g(.)$ determines a composite of effective labour in efficiency units, given inputs of family (F) and hired labour(H)

V represents other production inputs

Θ is the parameter set of the g (.) function which determines the nature and extent of labour heterogeneity

α is the parameter set of production function f (.) with aggregate labour and other inputs

u is the random error term.

The test for labour heterogeneity consists of specifying alternative functional specifications for the g(.) function and testing whether the parametric restrictions which lead to a specification of labour homogeneity are supported by the data. We test several functional forms for the $g(.)$ function which creates an aggregate labour composite from observed levels of family and hired labour, allowing for differences in marginal productivity as well as for a constant or varying AES between family and hired labour.

Labour heterogeneity has two related dimensions:

(i) Are family and hired labour perfect substitutes in the production function in the sense that the Allen partial elasticity of substitution (AES) between family and hired labour is infinitely large?

(ii) Are the marginal products per unit of labour time of these two types of labour equal to each other, everything else held constant?

---

[15] There is a prior step involved in which it is necessary to test whether F and H can indeed be represented in a separable composite given by the g (.) function. These tests for separability are described and carried out in Chapter VI.

Dimension (i) above is a test for whether the $g(.)$ function is linear. Dimension (ii) is a test for whether the marginal rate of substitution (MRS) between family and hired labour is always equal to one.

As discussed in Chapter III, when F and H can be aggregated in a separable composite given by the g (.) function, the relative efficiency of hired and family labour is determined solely by the parameter set Θ of the g (.) function. At the optimal labour allocation let the ratio of the marginal productivity of hired labour *vis a vis* family labour be represented by θ^*.
With separability,

$$(4.10) \qquad \frac{\partial Q / \partial H}{\partial Q / \partial F} = \frac{\partial g / \partial H}{\partial g / \partial F} = h(F^*, H^*, \Theta) = \theta^*$$

Consequently, the value of θ^* estimated from the production function can be used to derive the effective wage rates and the shadow profits required for the estimation of the labour supply component of the model in the second step. Instead of defining the shadow wage rate for family labour, w^*, in terms of the estimates of the marginal product of family labour from the production function, Equation 4.6 can be re-specified in terms of the observed market wage rates, w and w^h and θ^* as follows:[16]

$$(4.11) \qquad Ls = \ell\,(w^*(\theta^*), \pi^*(\theta^*) + E, \mathbf{S}, \beta, T) + e$$

$$\begin{array}{lllll} \text{where } w^*(\theta) & = w & \text{if } M > 0 & \text{(small farm)} \\ & = w^h/\theta^* & \text{if } H > 0 \text{ and } M = 0 & \text{(big farm)} \\ & = w^- & \text{if } H = M = 0 & \text{(autarchic farm)} \\ & & \text{where } w \leq w^- \leq w^h/\theta^* & \text{for } \theta \leq 1 \end{array}$$

$$\begin{array}{llll} \pi(\theta^*) & = Y^*_N - wF & \text{if } M > 0 & \text{(small farm)} \\ & = Y^*_N - (w^h/\theta^*)\,F & \text{if } H > 0 \text{ and } M = 0 & \text{(big farm)} \\ & = Y^*_N - w^-F & \text{if } H = M = 0 & \text{(autarchic farm)} \end{array}$$

---

[16] Assuming $w < w^h/\theta^*$. Otherwise all family labour is supplied on the off-farm wage market by all households. See the discussion in Section 3.5 of Chapter III.

where Y^*_N is the optimized value of net farm income including returns to family labour (Refer to Eq. 4.8c).

S = a vector of household and individual level characteristics which affect preferences between leisure and consumption

e = error term.

The advantage in the estimation structure of Equations 4.9 and 4.11 is that it avoids direct use of the estimated marginal product of family labour in the labour supply equations.[17] Predicted values of the marginal product of family labour at the household level is likely to show an extreme level of variation which may not be consistent with the variation in the unobservable underlying subjective evaluation of family labour, U_{ℓ}/U_C. Unless certain restrictive functional forms are used, the estimated marginal product of family labour could be negative in a large percentage of the sample households from which the production function is estimated. In Jacoby's own exercise when the production function was estimated as a fully specified translog equation, nearly 20% of the sample households produce a negative marginal product for female family labour.[18]

Another problem with using marginal products of labour in the labour supply equation to derive both the shadow wage and the shadow farm profits is that it increases the likelihood that errors of the production function and labour supply equations (u and e) will be correlated. Such correlations increase the complexity of the two-step estimation procedure and the adjustments in the standard errors in the second step which must be made in order to draw correct statistical inferences. (This adjustment process is discussed in Section 4.4 below). While the potential

[17] In the specification of Eq. 4.11 an estimate of the marginal product of family labour is required only to exactly identify the effective wage rate of the autarchic household since in equilibrium w^* will equal the marginal product of family labour. The alternative specification based on the observed market wage rates w and w^h and the θ^* value, however, can still be used to define a fairly narrow range over which the effective shadow wage rate must lie even in the case of the autarchic household.

[18] Jacoby addresses this problem by dropping specific interaction terms involving female family labour that have a negative coefficient in the fully specified translog equation with some loss of generality in the flexible functional form. (Jacoby 1993: 913 footnote 14).

endogeneity of the marginal product based measures of w* and E* in Eq. 4.11 could be addressed by using instruments for the shadow wage rate and shadow farm profits, the choice of instruments that are not only uncorrelated with e but also uncorrelated with u are likely to be limited.

The two step sequential estimation strategy proposed above - where some of the regressors used in the labour supply equation are derived from parameters estimated in the production function equation - is an example of what Pagan (1984) refers to as a two step procedure for estimating models with a "generated regressor". In these models the main interest is in obtaining consistent estimates of the parameters in the second stage regression when it contains variables that, while directly unobservable, are estimable from the parameters of an auxiliary equation estimated in the first step.[19] Our interests, however, are somewhat different from the usual model with generated regressors. The auxiliary equation of the production function does not serve only to generate the unobservable variables for the labour supply regressions but also to detect the presence and nature of the heterogeneity between family and hired labour as reflected in the Θ parameter set.

An alternative strategy to the two-step procedure would be to estimate the first (auxiliary) and second step models via some joint method, such as full information maximum likelihood (FIML) as in Liederman (1980). In reference to Eq. 4.9 and 4.11, FIML estimation would imply that the Θ parameter vector reflecting the heterogeneity between family and hired labour would be estimated jointly from the production function and the labour supply equations. Our tests for labour heterogeneity would likely be more robust if the estimates of the Θ parameter set were derived from a joint estimation of the production and the labour supply function. In addition there would be the usual efficiency gains if the error terms in the two equations were to be correlated. A joint estimation strategy for Eq. 4.9 and

[19] In econometric applications models in which generated regressors occur are widely used. Some example are models of (rational) expectations where the variable reflecting the anticipated value or expected value is generated as the predictors from another equation representing the expectations process (Barro 1977; Topel 1982); and spatial models of labour market equilibrium which include as a regressor the predicted probabilities for unemployment (Abowd and Ashenfelter 1981).

Eq. 4.11, however, is equivalent to joint estimation of the production and consumption components of the underlying farm household model. As noted in Section 4.1, such a joint estimation strategy for farm household models is computationally complex; and, depending on the nature of the heterogeneity implied by the Θ parameter set, it may not always be feasible to solve analytically for the reduced form equations. If so, one cannot take advantage of economic theory in imposing or testing for parametric restrictions in the estimated equations.[20]

Another reason why FIML is not used in this study is the structure of the data. The farm production function is estimated at the level of a household while the labour supply regressions are done with individual household member data. With such a data structure it is awkward to hypothesize the joint distribution for the error structure in the first and second step regressions. While the labour supply regressions could also be estimated at the household level by averaging over individual members, it is a better option to work with the individual level data since individual characteristics can be important in determining of labour supply behaviour in addition to shared household level characteristics.

While opting for the sequential estimation procedure in which the Θ parameter set is derived form the production function equation only, we carry out sensitivity analyses to show that the production function based results on labour heterogeneity are robust to many alternative assumptions and specifications of the production function (Chapter VI). Secondly, while the labour supply equations do not directly contribute to the estimates of Θ, they can be used to provide independent corroboration of the nature of labour heterogeneity detected in the production function estimation. This is accomplished through appropriate model selection diagnostic tests for comparing alternative model specifications (in Chapter VII) that allow for a common effective wage (which is consistent with family and hired labour being homogeneous inputs) and varying effective wages conditional on the estimated Θ (which is consistent with labour heterogeneity).

---

[20] Singh, Squire and Strauss (1986c: 21).

4.4 Error Correction for the Two Step Estimator

The basic strategy of the two step procedure is to replace the unobserved regressors in the second step with their estimated or predicted values from the auxiliary or first step regression. These values are then treated as if they were known *a priori* for the purpose of estimation and inference in the second step model. That the two -step procedures yield consistent estimates of the second step parameters under fairly general conditions is well documented as is the fact that the second step standard errors and related test statistics reported in the normal regression output are normally incorrect (Pagan 1986, Murphy and Topel 1985). But the need to correct the standard errors in the second step is commonly ignored.

The standard errors in the second step require an upward adjustment in order to account for the fact that the generated regressors are measured with sampling error which is related to the precision of the parameter estimates in the first step. In an illustrative example reported in Murphy and Topel (1985: 372), the proportional adjustment in the estimated standard errors is largest for the generated regressors. The required adjustments are somewhat smaller for variables that appear in both the first and second step equations, and negligible for the non-generated exogenous regressors in the second stage which do not also appear in the first step equation.

In the specific framework of the estimated model for this study the key parameters of interest in the second step labour supply equations are the wage and income elasticities of labour supply. Since the wage and (non-labour endowment) income variables are the generated regressors in Eq. 4.11, for which the required adjustments in the standard errors of the relevant parameters are likely to be the largest, it will be important to make the necessary adjustments.

The procedure adopted for adjusting the second step standard errors follows that suggested by Murphy and Topel (1985). These adjustments are relatively straightforward when the second stage equation is linear as in Equation 4.11 and when it is further assumed that the random errors of the first step and second step equations are un-correlated.[21] This latter simplifying assumption is also made in the estimation work of Jacoby (1993).

The general structure of a two step estimation procedure can be written as:

$$(4.12) \qquad q = x_1\alpha + u$$

$$(4.13) \qquad y = x_2\beta + f(\alpha, x_1)\gamma + e$$

where α is a vector of unknown parameters estimated in the first step, based on a vector x_1 of exogenous variables which determine q. And y is the variable of interest in the second step regression and y is influenced both by an exogenous set of variables, x_2, and another set of variables which are actually unobservable but which can be generated with the help of the α parameters (and the x_1 variables) of the first step equation. These unobservable variables are given by $f(\alpha, x_1)$. The set x_1 and x_2 could, of course, overlap and f(.) need not be related in any way to the function which determines q from x_1. The only restriction imposed on f(.) is that it be twice continuously differentiable for each α in x_1. Hence f(.) could be non-linear in α.

It is assumed that the first step regression yields an estimator $\hat{\alpha}$ of α which is consistent and asymptotically normal with covariance matrix $V(\hat{\alpha})$. Let $\hat{V}(\hat{\alpha})$ be a consistent estimate of $V(\hat{\alpha})$ which is obtained from the first step regression. In the second step y is regressed on x_2 and the estimated values $f(\hat{\alpha}, x_1)$ and what is required is the correct asymptotic distribution of (β, γ) for tests of statistical significance and other inference based on a least squares regression of Eq. 4.13.

Let r be the number of elements in the $\hat{\alpha}$ vector and n the number of observations for the second stage regression. Denote by X_2 be the n x p matrix of observations on x_2 and let B be the n x m matrix of generated regressors $f(\hat{\alpha}, x_1)$. Denote the full observation matrix of the second step equation as $Z = (X_2, B)$. The covariance matrix for (β, γ) when Eq. 4.13 is estimated by ordinary least squares is given by:

[21] Under these conditions the standard error adjustments can be made on the basis of the normal regression output of most econometric packages. When the errors are correlated the adjustments are more complicated. See Newey (1984) for details.

$$\sum_0 = s^2 (Z'Z)^{-1} \tag{4.13}$$

where s^2 is the sample estimate of the variance of e.

The correct asymptotic covariance matrix for (β, γ), adjusted for the fact that $\hat{\alpha}$ is estimated with error, is given by[22]

$$\sum = \sum_0 + (Z'Z)^{-1} Z'B^* \hat{V}(\hat{\alpha}) B^{*\prime}Z(Z'Z)^{-1}. \tag{4.14}$$

where B* is computed from the derivative matrix of f(.) with respect to α such that the typical element b^*_{ij} of B* is given by [23]

$$b^*_{ij} = \sum_{k=1}^{m} \gamma_k \frac{\partial f_k(\hat{\alpha}, x_{1i})}{\alpha_j}. \tag{4.15}$$

The form of Eq. 4.14 indicates the error correction procedure for the second step equation exceeds the commonly reported covariance matrix $\sum_0$ by a positive definite matrix (given by the second term in Equation 4.14). As a result the standard errors based on the unadjusted $\sum_0$ in a naïve two step procedure are always under-stated. It is also clear that the size of the adjustment in the standard errors of (β, γ) depend crucially on two factors: $\hat{V}(\hat{\alpha})$ which is the precision of the first step parameter estimates, and the correlation between the explanatory variables Z and the derivative matrix B*. To the extent that this correlation is high, or the first step sampling error significant, the error adjustment will be important (Murphy and Topel 1985:375).

4.5 Summary

The estimation methodology for this study follows a two step procedure adapted from the method proposed by Jacoby (1993) for estimating non-recursive farm household models. In the first step, an aggregate farm production function

[22] Murphy and Topel (1985: 374-375).

[23] In Equation 4.15 the index i represents the observation number (from 1 to n), the index j ranges over the number of elements in the α parameter set (from 1 to r); and k indexes the number of generated regressors in Eq. 4.12 given by f(.), which by assumption is 1 to m. Hence the matrix B* is n x r while B is n x m.

containing a separable labour nest is estimated, enabling the heterogeneity between family and hired labour to be tested. Several alternative specifications that allow for imperfect substitution between family and hired labour and varying marginal products are estimated. In the second step a structural labour supply equation is estimated with a specification that is consistent with the type of labour heterogeneity found in the production function estimation. This typically means that the effective wage and non-labour income variables used in the labour supply regression will be derived from parameters estimated in the production function.

Jacoby's approach was to base these "generated regressors" in the labour supply equation on the marginal product of family labour estimated from the production function. In Chapter III we showed that the marginal product for family labour at the optimum labour allocation could be directly related to the observed market wage rates for family and hired labour, when the heterogeneity between family and hired labour is explicitly modeled within a separable labour nest in the production function. Depending on the net labour market position of specific households, it is possible to derive the effective wage rate that determines their labour supply equilibrium from the observed market wage rates and the parameters that reflect the labour heterogeneity. This provides an alternative specification of the labour supply equation that has several advantages over the specification based on directly using the estimates of the marginal product of family labour.

An alternative estimation strategy based on full information methods to estimate the production and labour supply equations jointly would have allowed for the nature of the labour heterogeneity to be simultaneously determined from the production and labour supply behaviour of the farm households. The complexities of such joint estimation, however, are formidable. The method adopted for this study is the simpler but widely used alternative of sequential estimation which gives consistent estimates under quite general conditions, requiring only a straightforward adjustment to the standard errors of the second step estimates.

CHAPTER V

THE SETTING AND THE DATA

This Chapter contains three sections. Section 5.1 contains a cursory description of the *tarai* (southern lowland) region of Nepal that provides the data used in the farm-household model estimation work of this thesis. It also briefly explains why the question of heterogeneity between family and hired labour in Nepalese agriculture is particularly relevant in the context of the *tarai* region. Section 5.2 describes the survey design and main features of the household survey and the structure of the data collected. The last section (5.3) provides additional information on the definitions adopted for the main variables used in the production function and labour supply regressions, and further details of how they were computed from the data recorded in the household survey.

5.1 The Setting

Nepal has three distinct ecological regions that lie on a north to south axis throughout the whole country. They are the northern most mountainous region, the middle hill region and the southern lowland *(tarai)* region. The material conditions for agricultural production are very different in these three regions. These differences derive mainly from climatic factors related to variations in altitude, but over the years such differences in physical conditions have led to very distinct regional agrarian structures and the social relations of production.

The empirical component of this thesis is based on household level survey data from the *tarai* region of Nepal only. The data are drawn from a large nationally representative household budget survey carried out by Nepal Rastra Bank (the central bank of Nepal) which is referred to as the Multi-purpose Household Budget Survey (MPHBS). Although the MPHBS collected household data from a representative sample for each of the three ecological regions of Nepal, this complete data set is not utilized in the present study. The estimation of the farm household model is limited to the data from the *tarai* sample because the question of heterogeneity between family and hired labour in Nepalese agriculture is best addressed in the context of that region

The MPHBS sample of households from the northern hill and mountain districts of Nepal has been excluded from this study because of the very limited use of hired labour in those regions. The average farm size in these northern regions is only about 0.7 hectares per operated holding (HMG/CBS 1993). Farm production is mainly subsistence oriented. There is only a limited form of economic differentiation among farm households in this region. Almost all households are owner-cultivators.

The agrarian structure in the southern plain (*tarai*) region of Nepal more closely follows the classical structure of a few large landlords dependent primarily on hired labour, a middle group of owner-cultivators and a large group of landless households who supply the hired labour on the larger farms. Although average farm size in the *tarai* region is still small (about 1.2 hectares), a wider socio-economic differentiation of the rural households exists. A significant percentage of the land is cultivated in operated holdings in excess of 10 hectares; while on the other extreme a significant proportion of the population consists of landless households who work as hired labourers on the big farms.

The National Sample Census of Agriculture for Nepal gives data on the size distribution of land holdings based on operated area and not on owned area. Hence there are few completely landless households enumerated in the census since those who do not own any land will usually rent in small operated areas. On the basis of the operated size of farm holdings from the 1991/92 Census[1], more than 10% of the *tarai* region households operate less than 0.1 hectares of land. Another 8% operate between 0.1 and 0.2 hectares. On the other hand, while holdings greater than 10 hectares comprise only 0.5% of the total holdings (households), but they account for about 7% of the total operated land.

[1] The information on the size distribution of land holdings in this section is compiled from the *1991/92 Nepal Agricultural Census* volumes (HMG/N, CBS 1993, Table 3).

The rural labour markets in the *tarai* region are quite active. In peak periods there is considerable in-migration of agricultural labourers from India as well (Wallace, 1989). Because of better regional transportation infrastructure, local villagers also have easier access to off farm wage employment opportunities in other villages and nearby urban centres. Rural households in the *tarai* region of Nepal are less likely to be quantitatively constrained on either the labour demand or labour supply side than farm households in the northern hill/mountain regions.

A distinctive feature of hill region farming is that households arrange reciprocal labour exchange arrangements among household groups of equal status. The incentive structures for exchange labour will differ from that for hired as well as own family labour. This is another reason why this study is focussed on the *tarai* region only, where exchange labour arrangements are very rare.

Nepal's *tarai* region is an extension of the Indo-Gangetic plains of northern India. Within Nepal the *tarai* comprises a narrow strip of land along the southern border at relatively low altitudes in comparison to the hilly north. *Tarai* region crop production is undertaken on level land, consisting of alluvial soil, which can be irrigated with the monsoonal flow of the region's rivers. Farms in the northern areas consist mainly of terraced hillsides that are difficult to irrigate.

The *tarai* region has about half of the population of Nepal and two-thirds of the cultivated land. The main summer crop is paddy, sometime grown in two rotations where irrigation is available. Wheat, pulses and oilseeds are the main winter crops. Maize and millet are the common rotation in upland plots (*pakho*) not suitable for paddy cultivation.

Agricultural production technology in the *tarai* region of Nepal is very backward. Average paddy yields are only about 2 metric tons per hectare and have been declining (HMG/N, MOF 1996) as more and more marginal land is brought under cultivation. Human labour is the main farm input, typically accounting for more than

50% of the total cost of farm cultivation.[2] There is a significant east-west sub-regional variation within the *tarai* belt of Nepal. The monsoon rainfall decreases as one travels west and hence so does population density.

The sub-regional economies of the *tarai* region are more closely integrated with the Indian border markets than with each other (Wallace 1989). Hence one observes considerable variation in output and input prices within the *tarai* region of Nepal. This variation extends to market wage rates for agricultural labour. This variation in the sample data makes it feasible to estimate the labour supply regressions with cross-sectional data from within the *tarai* region only.

5.2. The Data Set

(a) Source

The Multi-purpose Household Budget Survey (MPHBS), was conducted in a phase-wise manner in 1984 and 1985 and it collected household level data with reference to the 1984/1985 annual cropping cycle. This survey is the most extensive of the periodic but irregular household budget surveys conducted by Nepal Rastra Bank (henceforth, Rastra Bank), designed principally to compile information for the preparation of consumer price indices. The MPHBS, however, had a more ambitious objective of improving the Nepalese database on the regional patterns of employment, income distribution and consumption because regular surveys of this nature are not conducted by other government agencies in Nepal.

The full sample size for this survey was 4022 households, of which the rural component consisted of 3660 households, drawn from a stratified sample of 22 districts (out of a total of 75 in Nepal). The data used for this thesis consists of a subset of the rural sample from the southern *tarai* region of Nepal. The full *tarai* region sample in the MPHBS was 1571 households from seven sample districts. The actual estimation work reported in Chapters VI and VII is based on a smaller sub-set

[2] See HMG/N, Ministry of Agriculture, *Costs of Production for Major Crops in Nepal 1985/86.* The data from this publication is contemporaneous with the MPHBS sample data.

of the *tarai* region sample, consisting of 1007 households. This is the complete household sample from five out of the seven *tarai* districts of the MPHBS.

The reason for the further reduction of the *tarai* sample size is that not all components of the MPHBS data used in this research was made available from Rastra Bank in computerized form. Because of the vast amount of data collected and the extensive work required for processing and data cleaning, some parts of the MPHBS questionnaire were not computer processed at Rastra Bank.[3] Unfortunately, one component of the survey left out of the Rastra Bank data tapes was the extensive data collected on the farm management module of the survey with the details of the crop-specific farm outputs and production inputs.

The data processing of the farm management component of the MPHBS was organized by the author of this thesis in 1992 at the Institute for Integrated Development Studies (IIDS), Kathmandu, from the original survey questionnaires. At IIDS it was not feasible to process the complete farm management component from all of the 3660 rural household questionnaires. An arbitrary choice was made to process the full household sample from only 13 of the 23 sample districts, based on regional representativeness. Although the additional data processing at IIDS was initially done for other purposes than the research work undertaken in this thesis, the coverage of the *tarai* sample used for this thesis is close to two-thirds for households (1007 out of a possible total of 1571 sample households), and even higher in the case of sample districts (5 out of 7).[4]

The production function estimates reported in Chapter VI are based primarily on the farm management data sub-set of the MPHBS that was processed at IIDS. This component of the survey has been linked with the computerized data files obtained

---

[3] The data processing and analyses of the MPHBS by Nepal Rastra Bank was considerably delayed in any case. The survey report was issued only in 1989 (Nepal Rastra Bank, 1989); and public access to the data set was provided only after the survey report was published.

[4] The five districts are Morang, Mahottari, Rupandehi , Banke and Kailali. Each of these districts represents one the five regional development zones of Nepal, moving from east to west, respectively. The other two *tarai* districts in the MPHBS sample were Sarlahi and Bara.

from Rastra Bank on other parts of the survey to prepare a full integrated data base of the 1007 sample households for the labour supply regressions of Chapter VII.[5]

In spite of the long period since the data were collected, the MPHBS remains one of the most detailed household level survey data collected for Nepal.[6] It was well designed and implemented, and the data collection was carefully monitored. Information was collected over several rounds. One of its special features was to collect data on many non-monetized transactions (e.g., income received in kind as gifts and transfers, or as a part of wages) which are important sources of income in poor developing countries but are often overlooked in household surveys.[7] In spite of its richness, the data from the MPHBS has been rarely used for subsequent research work on Nepal. This neglect provided part of the motivation to base the estimation work of this thesis on the MPBHS data.

(b) Sample design

The sample design of the MPHBS was based on a four-step selection procedure. In the first step 22 sample districts were chosen based on random sampling from a regional stratification of all the districts of Nepal.[8] Within each sample district the subsequent steps followed in selecting the sample of rural households consisted of:

---

[5] In hindsight even the limited coverage of the MPHBS sample attempted in the IIDS data processing proved overly ambitious. The computerized data processing, cleaning and, most importantly, the linking of the farm component module with the data set received from Rastra Bank proved more difficult than anticipated. These problems considerably delayed the research work carried out for this thesis.

[6] A Poverty Assessment report on Nepal was prepared by the World Bank based on the MPHBS data set (World Bank 1991) and the Bank has remarked favourably on the quality of the data in this survey. Subsequently, the World Bank itself conducted a Living Standards Measurement Survey (LSMS) of Nepal in 1996 but access to the LSMS survey data was provided too late to be used in the estimation work of this thesis. In any case, the farm management component of the LSMS is not as detailed as the MPHBS.

[7] The data on non-monetized transactions and other own-household accounts proved useful in deriving the non-labour endowment income, especially for the poorer households, which is required for the labour supply regressions.

[8] The procedures followed in this step are not relevant for the purpose of this study which is based on an *ad hoc* sub-sample of the *tarai* districts selected for the MPHBS. The details about the regional stratification of the sample districts and selection procedure for the MPHBS are given in NRB (1988) Chapter 2.

(ii) a random selection of sample village *panchayats*

(iii) a random selection of 2 out of 3 clusters[9] from each sample *panchayat*

(iv) a stratified sample of households from each cluster derived from five strata based on the size of the land holding by the household.

The *panchayats* selected in step (ii) correspond to the main administrative units within a district.[10] The *panchayat* boundaries in Nepal are drawn to include about 1,000 to 1,200 households, and so they do not usually coincide with a single village. Nonetheless, the geographical area covered within a *panchayat* boundary tends to be homogenous, particularly in the *tarai* region which has a much higher population density than the northern hill regions of Nepal. Hence, for convenience of language, in the subsequent text of this Chapter and elsewhere in the thesis, the term "village" is used interchangeably with village *panchayat*.

In the last step of the sample design there was a complete enumeration of all households within the selected clusters. These households were then classified into five sampling strata (categories) based on the size of the operated land of the farm household. Operated land consists of owner-cultivated land and land rented in. The first four strata were defined to be households with operated farm sizes designated as marginal, small, medium or large farm holdings.[11] A fifth stratum consisted of landless households. Within each stratum a five percent random sample was drawn, with an upward adjustment in the sampling ratio where necessary to ensure that at least two households were selected from each stratum.

---

[9] Each village *panchayat* is sub-divided further into 9 wards. In the MPHBS sample design each selected *panchayat* was broken down into 3 clusters, each consisting of 3 adjoining wards.

[10] The village *panchayats* in Nepal have been renamed as village development committee areas after the political changes of 1991 but the boundaries have remained the same.

[11] The farm size strata were defined in terms of the area of total land operated by a household. Operated land includes self-cultivation of owned land as well as land rented in. So the distribution of operated land will differ somewhat from and tend to be more equal than the distribution of land ownership. The size strata limits were as follows: marginal farm holdings up to 1.02 hectares, small holdings up to 2.73 hectares, medium holdings up to 5.44 hectares, and large farm holdings above 5.44 hectares. Households with only a homestead that may have included a kitchen garden were classified as landless.

The important implication of this sample design is that within each sample village *panchayat* the material conditions of farm production, the cropping pattern and wage rates and prices will tend to be fairly uniform. But these will vary considerably across *panchayat* and districts, especially since the five district sub-sample has been selected from the entire range of the east-west regional variation in Nepal. Within a village (and consequently within ward clusters) the main basis for the variation in the household level data will be with respect to income and assets holdings based on the amount of land cultivated. Hence the MPHBS data is a broadly representative sample from different socio-economic strata within a village. Such a variation proves useful in the labour supply estimation work to identify the income effects across households that face more or less equal wage rates.

The final sample of 1007 *tarai* region households used in the empirical part of this thesis does not constitute a completely random sample for the *tarai* region of Nepal. The sample of households within districts is the complete set selected for the MPHBS, and hence can be treated as random (subject to the stratified clustered sample design described above). But the random sample properties of the overall *tarai* sample of the MPHBS do not carry over to the data subset drawn from only 5 of the 7 *tarai* districts of the MPBHS. In the regression estimations of this thesis no specific attempt is made to deal with this problem by modifying the weights given to the households to reflect the *ad hoc* selection of the 5 districts. The focus of this study in not on estimating *tarai* region aggregates or household means. The estimation work assigns equal weight to each household. To the extent that the full sample of 1007 household is used, the equal weight procedure is justified since the sample size within districts in the MPHBS was chosen to be proportional to the district population.[12]

The distribution of the actual sample of households used in this thesis, classified by sample district and land size strata, is given in Appendix Table 5A.1.

---

[12] At the final stage of the stratified sample design of the MPHBS there is also a more or less equal representation of households from the five land size strata (NRB1988: 29). This property is maintained in the five district sub-sample since the entire sample within a district is used.

(c) Data structure: farm management

The farm management module of the MPBHS collected detailed data on all output and input use on the farm over a multi-round twelve month survey period. The farm household's own-account production activities were classified into agricultural and non-agricultural enterprises. The main sub-categories identified under agricultural enterprises were crop production, animal husbandry, horticulture and fishery. Crop production is the dominant farm activity. For the full *tarai* region sample of the MPHBS the crop production component constitutes more than 81% of total agricultural enterprise income; and the latter (own account agricultural enterprise income) itself accounts for about 65% of total income for the average *tarai* region rural household.[13]

The MPHBS collected very detailed output and input data on crop production at the level of individual crops. In particular the labour input in specific crops was classified into nine categories based on three types of workers (adult male, adult female and child) and three sources (family, hired and exchange labour). Detailed information was collected on the actual wage payments and meal expenses related to hired and exchange labour so that an average wage payment for each type of labourer can be computed for each household reporting use of hired and exchange labour on a crop-wise basis. Data on other production inputs such as bullock days and seeds are also distinguished by source - whether hired (or purchased) or from the household's own resources. In almost all categories the information is collected in both quantity and value terms with the main exception being chemical fertilizer.[14]

[13] NRB (1988:230) Table 2A. Other sources of household income are income from wages and salaries, rental income, non-agricultural enterprise income (i.e. household trade and cottage industry activities) and gifts and remittances. Within the MPHBS definitions of own account agricultural enterprise income, apart from crop production which accounts for 81.2% of this sub-total, the minor components are animal husbandry (16.1%), horticulture (1.4%) poultry (1%) and fisheries (0.2%)

[14] The absence of quantity data on chemical fertilizers is not a serious shortcoming in the estimation of the production function. Fertilizer distribution in Nepal occurs through a single para-statal agency which charges a fixed price throughout Nepal. So the value data is a good proxy for quantity if one ignores differences in the nutrient content of different fertilizers.

The farm management data set is organized into a winter and summer cropping cycle/rotation. This structure can be used to derive average wage rates paid for hired labour for crop production in each of the two cropping cycles to get a measure of seasonal variation in the wage rate.

Although the farm management survey data, as with the rest of the MPHBS data, was collected through recall based interview methods, several steps were taken to ensure reliability and accuracy of the farm management data. For instance, the data on crop output was cross-tabulated with a table on the disposition of farm output (market sales, home consumption, remaining stocks) to produce a household balance by crop. Similarly, the information on family labour input in crop production was cross-checked with the labour supply component of the survey questionnaire to maintain consistency.

Detailed data were also collected on the size and type of land operated by each sample household. The farm-level land holding data is also distinguished by land tenure forms: owner cultivated, leased out and leased in either as fixed rent or share cropping arrangements. The main distinction in agricultural land in the *tarai* region is between wet land (*khet*) suitable for growing paddy and sloping upland (*pakho*) that does not retain water for paddy cultivation. Within each category a further distinction is made between irrigated and un-irrigated land. This makes it relatively easy to control for the differences in the quality of the land input. The cropping pattern between *khet* and *pakho* land tends to be very different. Hence defining the land input to take account of the heterogeneity between *pakho* and *khet* land is one way to account for differences in the crop composition across farm households. While this information gives a good indication of the quality of the land input available to each farm household, the area planted to each crop, unfortunately, is not broken down into these specific land types. The crop-level land input is recorded as a single value for the total area harvested.

(d) Data structure: Family labour supply

The family labour supply component of the MPBHS data is also very detailed. A complete enumeration of all household members with details on individual characteristics (age, marital status, education level, etc.) was prepared in the first survey round. Members of the household who are part of the extended farm family are clearly distinguished from temporary residents and farm servants. The usual occupation for each household member is noted on the basis of a very detailed three digit coding structure, and this information is used to determine whether a particular person is economically active or not. For each economically active person (aged 10 or above) data on the seasonal employment pattern was collected through multi-round interviews. These data were then aggregated into total days of productive work in each month of the survey year. The main categories of employment distinguished for the number of days worked in each month are: work on own family farm, work on the hired labour market and work as an exchange labourer.

A major shortcoming of the labour supply data set is the lack of information on the wage rate received by each individual when working in the hired labour market. [15] Nor does the data distinguish the place and type of employment - i.e., whether it is agricultural labour or non-agricultural employment, and whether it is within the sample village or outside. This deficiency means the wage rates for individual family members who report working on the off-farm labour market must be imputed from some other part of the MPHBS data set.

5.3 Main Variable Definitions

(a) Production function

The production function estimations reported in Chapter VI are based only on the crop production component of the farm management data which is the dominant part

[15] There are data on monthly income from off-farm wage labour at the household level. While this could be used to derive an average earning per off-farm work day at the household level, this computation is likely to be affected by major reporting errors. Another problem is that this information cannot be used to derive wage rates for male and female workers separately.

of farm production in the *tarai* region. The other components of agricultural production identified in the MPHBS - animal husbandry, horticulture and fisheries - are ignored. These other activities are ignored because they are minor components of farm production;[16] and because detailed input data are not always available for these other components in the MPHBS. Finally, there is very little use of hired labour in these ancillary activities and so they are not relevant to an estimation framework where the main objective is to test for the heterogeneity between hired and family labour.

Although the MPHBS contains data on crop-specific farm inputs and output, the test for labour heterogeneity is embedded in an *aggregate* production function for all crop output mainly because the important information on land quality is not available at the level of individual crops. The differences in the efficiency of family and hired labour may also lead households to choose alternative cropping patterns based on the labour requirements of different crops. One particular manifestation of the labour efficiency related production choices made by households may be more intensive multiple cropping. Such responses will not be fully captured in comparisons of labour productivity in single crop production functions. Further, since the family labour supply data are also at the aggregate level (i.e. not distinguished by crop specific work days) it makes sense to derive the effective wage rate for family labour in terms of an average comparison of the productivity of family and hired labour, rather than crop-specific comparisons which may vary considerably among crops.

On the other hand, working with aggregate farm output does create potential bias in the estimation results when the crop composition and output prices are not uniform across different households and regions (Bardhan, 1973). The price variation problem can be accounted for by measuring the farm output as a composite quantity

[16] As noted in footnote 13 on page 102, the role of horticulture, poultry and fisheries components in total agricultural enterprise income is negligible for the average *tarai* region rural household. While animal husbandry, which on average accounts for 16% of total agricultural enterprise income, is an important ancillary activity to crop production, there is almost no use of hired labour for this component. Also, it is difficult to correctly impute what annual animal husbandry production and income is when sales and slaughter of farm animals also represent a depletion of capital stock.

variable rather than simply as the market value of total output. The quantity index for aggregate crop output is computed by deflating the total value of farm output by a village-specific aggregate price index. The village specific price indices are derived as a Tornquist (log-linear) price index of individual crop prices, using the value share of different crops grown in that particular village.

From Eq. 4.9 in Chapter IV, the specification of the production function with a separable nest for labour inputs is of the form:

$$(5.1) \qquad Q = f(\, g\,(F, H, \Theta), V, \alpha) \; + \; u$$

where Q is aggregate composite output as derived above.

On the input side, in addition to the land (A) and composite labour input $g(.)$, two other main input variables were created from the MPHBS survey data (to represent the set V). They are bullock-pair work days (B) and material inputs (M).

The material input variable simply lumps together several diverse production inputs identified in the MPBHS data in value terms. M is the sum of the value of seeds, chemical fertilizers, insecticides, irrigation charges, and operating cost and rental charges for farm machinery. The bulk of M consists of the value of seeds and fertilizers since only a small percentage of farms reported machinery use or payment of fees for irrigation water. M is converted into real units by deflating the nominal sum by the village-specific crop price index. The use of such a deflator is justified since the major component of M is the value of seeds.

Apart from the components of M, all of the other major inputs are already reported in quantity levels in the MPBHS data. The main computations required for defining these variables are aggregation procedures over sub-categories.

As noted above, for the labour inputs, the MPHBS reports work days on the family farm by male adult, female adult and child labour under each of three sources of labour: family, hired and exchange. The input of child labour is not broken down further by gender. Fortunately, the overall level of child labour reported as inputs in crop production is minimal in the MPHBS farm management data in the selected *tarai* districts.[17] Hence the child labour category is completely ignored in creating the sub-aggregates of total family and hired labour. The level of exchange labour in the *tarai* region is also minimal. Where it occurs it has been treated as hired labour.[18]

The adult male and female work days in each category are added together after imply adjusting the female work days into equivalent male labour days based on the observed male female wage rate in a particular village. That is, total family labour days (F) is measured in terms of adult male equivalent units and is computed as $F = F_m + (w_f/w_m)\ F_f$ where the *f* and *m* subscripts signify male and female sub-categories. A similar procedure is used for aggregating hired labour.

This specification treats male and female labour as perfect substitutes, adjusted for a constant productivity difference represented by (w_f/w_m) -- the observed ratio of the female to the male wage rates in a sample village. This is admittedly a crude manner of handling any heterogeneity between male and female labour (within both the hired and family labour sub-aggregates). It is, however, theoretically justifiable to the extent that differences in competitively set market wage rates will reflect differences in labour productivity at the margin.

Alternative treatment of the potential heterogeneity of male and female labour was attempted by distinguishing four different labour inputs in the labour aggregator *g*(.) function. Using such a structure with the flexible functional specifications of the production function proved intractable. Instead of persisting with an awkward four

[17] The MPHBS Report indicates that child labour is concentrated more on what it classifies as household subsistence activities - collecting drinking water and firewood, home processing of food crop, etc. (NRB 1988:141 Table I).

[18] Labour exchange arrangements can be viewed as hired labour contracts where payment is received in kind terms of the other participants' labour input. Of course, the incentive structure and costs of monitoring the effort of exchange labour differ from that of purely casual hired labour. In some specific situations the exchange labour arrangements may be based on close-knit groups or extended family/clan identities and hence be more similar to family labour. Since exchange labour work days in the MPHBS data for the *tarai* region are so limited, its classification ultimately is not important.

input $g(.)$ function this study has opted for the alternative approach of carrying out detailed sensitivity analyses to show that any observed heterogeneity between family and hired labour is independent of the specific values of the (w_f/w_m) ratio used to convert female labour work days to equivalent male work days.

The bullock work-days variable is also distinguished by family source and hired source. Our derivation of the total bullock input variable (B) is simply the sum of hired bullock days and family bullock days. Potential heterogeneity in input of bullock power from family and hired sources is not of interest in this study.

The land input variable however required some detailed computations. From the MPHBS farm management data it is possible to compute several alternative indicators of the land input. Since the dependant variable is aggregate farm output, the main distinction is between net sown area - a measure of the physical size of the land area that is cultivated by the farm household - and gross harvested area which takes into account of the actual cropping cycle whereby more than one crop may be planted on a particular plot in the reference period. In empirical applications production function specifications have used both types of the land input variable. Where data are available, gross cultivated area seems a more appropriate measure of the annual flow of services received from the physical land endowment of the farm (Carter 1984).

The problem with defining the land input in terms of the gross cultivated area with the MPBHS data set is that the information about land quality (paddy land or upland, irrigated or unirrigated) is available only at the level of the physical land endowment and not in terms of the land area allocated to specific crops. Several alternative ways of bringing in the land quality variables were tried. In the production function specifications using ordinary least squares estimation, additional land quality variables can be introduced in a ratio format - such as the ratio of irrigated land to total physical land or the ratio of paddy land in total land - by assuming that these ratios carry over to the gross cultivated area as well. Alternatively, in the specifications based on non-linear estimation, an aggregate quality adjusted land input can be computed directly from the physical land type data, and the index of cropping intensity in the same way as the aggregate labour input is constructed as a composite of family and hired labour.

(b) Labour supply function

From Eq. 4.11 the labour supply regressions are of the form

$$(5.2) \qquad L_s = \ell(\, w^*(\theta^*),\, E^*(\theta^*),\, S) + e$$
$$= \ell(\, w^*(\theta^*),\; \pi(\theta^*) + E\,,\, S) + e$$

where L_s is a measure of the total labour supply in all productive activities.

The labour time of family farm household members has four general uses:

(i) unpaid work on own farm cultivation, including time devoted to subsidiary productive activities such as livestock rearing, home processing of farm output for sale; etc.

(ii) work as a hired labourer in other households' farming activities

(iii) off-farm non-agricultural hired labour work (such as in construction work)

(iv) household subsistence activities, such as home processing for own consumption, time spent on collecting firewood and water, etc.

The definition of labour supply adopted (based on the definition followed in the coding of the MPHBS data) includes items (i) to (iii) only. [19] In the setting of the *tarai* region of Nepal there is very little off-farm non-agricultural work available for individuals who continue to reside on their farm. The bulk of the labour days of work reported in the MPBHS sample therefore consists of items (i) and (ii).

The labour supply regressions are run at the level of individual family members of the farm household in order to allow for the effects of variation in individual

---

[19] When data exists on activities included in (iv), this category can be treated as much a part of labour supply as the unpaid farm work under (i). Under optimal labour allocation conditions, the marginal returns to labour in all activities are equalized. Jacoby (1993) specifies his labour supply variable to include category (iv) activities.

characteristics such as age and educational levels. The sample of individuals is limited to immediate family members residing in the household who can be identified through their relationships to the household head.

The dependant variable is the number of days of productive work reported by each economically active family member. For each economically active family member the data on the number of days worked is available on a month to month basis. The monthly data can be aggregated into two six-monthly sub-totals to correspond to the two phases of the survey data collection in the farm management module. This procedure helps to establish a link between the labour supply work days (in six monthly sub-totals), and data on the average wage rates paid by farm households that report hiring in labour in the summer and winter cropping cycles. In this way the seasonal labour supply behaviour of a particular individual can be modeled since there will be differences in the average wage rates reported in the two cropping cycles. It is at best a gross simplification to assume that the operation of the rural labour markets in the study areas gives rise to two distinct seasonal wage rates, each of which is constant over an arbitrary six month survey cycle. Nonetheless, this method of aggregating the monthly labour supply data offers a crude way of accounting for some seasonal wage variation to compensate for the missing data on the actual seasonal wage rates received by individual household members.

As noted above, the main definitional treatment regarding the wage rate variable is the imputed equivalence of the off-farm wage rate for family labour and the wage rate at which labour is hired in by the sample households in a particular village. In effect, the assumption $w = w^h$ is forced since w is not directly reported in the MPHBS data set.[20] Although such an imputation for the off-farm wage rate for family labour could be a source of unspecified bias in the estimates of parameters of

[20] Actually what is required is w_{ijk} where the i indexes an individual in farm household j for the month k. The wage rates received by individual family members who report working off-farm in the hired labour market may vary according to their skill and nature of the specific tasks performed, even in a particular local labour market. Unfortunately this information is not recorded in the MPHBS data nor can proxies for it be computed from other parts of the survey questionnaire.

the labour supply function, such errors seem unlikely to be serious in the specific setting of the Nepal *tarai*. There are several reasons for this presumption.

Given the fragmented nature of rural labour markets in Nepal, the primary place for off-own-farm employment of family labour will be as hired labourers in the larger farms in their own or nearby villages. Hence it is expected that $w = w^h$ on average for a sample of farm-households from a particular village with limited non-agricultural employment prospects.[21] In a similar setting to the Nepal *tarai*, Bardhan and Rudra (1981) report a very limited locus of off-farm employment in agricultural operations for wage labourers in West Bengal. If the local village economies were completely closed, and crop production was the only economic activity undertaken, there would be an exact equivalence between the average wage reported as paid out and the average wage rate reported as received, even in the presence of seasonal variations in the local village wage rates. In reality local village economies are not completely closed, and there are other wage labour activities apart form crop production. But the effects these have in creating a major wage gap between the wages paid out and wages received by households in a particular village is still likely to be small.

Secondly, the analytical results in Section 3.5 of Chapter III indicate that w is the correct wage rate to use in the labour supply regression for farm household members only if $w > w^h/\theta$ consistently in the sample data.[22] But in such situations there would be complete specialization in labour allocation: family labour would be completely supplied on the off-farm market at wage w, and farm cultivation would be done completely by hired labour paid a wage w^h. Since one rarely observes such a pattern of specialization, it is more usual to find $w < w^h/\theta$. When the latter condition holds the effective wage rate to be used in the labour supply equation is defined in terms of

[21] This will be particularly true for female family members whose off-farm mobility is even more restricted. The wages female labourers receive will be the wages other households in the village pay to female farm workers in that village.

[22] θ is the ratio of the marginal product of hired labour and family labour at the optimum labour allocation for crop production. Refer to the discussion in Section 3.5 of Chapter III.

w^h/θ for the subset of households who hire in some labour for their farm cultivation. In such cases the off-farm market wage rate, w, need not be known. The verification of whether $w < w^h/\theta$ can be obtained on the basis of the observed labour allocation pattern alone. The individuals who receive a high off-farm wage rate would not be specializing in farm production work, and hence can be detected through the usual occupation codes reported in the MPHBS.

To minimize the discrepancy related to varying wage rates due to differences in the type of work performed, irrespective of the relationship between w and w^h/θ, the labour supply regression in Chapter VII will be limited to individuals who report their main occupation to be farm operators or agricultural labourers. For instance, the reported wage rate for hired labour in crop production would not apply to the labour leisure equilibrium of a household member who may do some work on the family farm but whose main occupation was as a shop keeper or a village carpenter. In these other principal occupations, the marginal returns to labour are likely to be very different from the average wage paid to hired labour in crop production. But the variations in individual level wages are unlikely to be important in village settings for individuals who report farm cultivation or agricultural labour as their main occupation, when farming is carried in under traditional cultivation techniques with little mechanization. The main difference will be between male and female adult wage rates which is recorded in the MPHBS data. Finally, the usual distinction between gross and net wages in settings such as the Nepal *tarai* is not likely to be important. The fixed costs to finding work on the local market will be minimal and there are no income tax wedges to consider.

It is concluded from the above that the imputation $w = w^h$, which is forced by the data structure in the absence of direct observations of w, is defensible in the proposed estimation set up and in the setting of the Nepal *tarai*. Hence, in the labour supply regression results of Chapter VII, w^h becomes the effective wage for any family member reporting work days in the hired labour category for a particular cropping cycle. For individuals in autarchic households, or in big farm households that hired in labour, the effective wage rates will be based on w^h with the appropriate adjustment for θ if labour heterogeneity is indicated.

The second main variable that needs to be created for the labour supply regressions is the measure of household non-labour income, including the imputed value of farm profits. From Eq. 4.7, the household-level nominal non-labour income (NLY), corresponding to the linearized budget constraint, is defined as follows:

$$(5.3) \qquad \begin{aligned} NLY &= E^* = \pi^*(w^*) + E \\ &= Y^*_N - w^* F^* + E \\ &\cong Y_N - w^* F + E \end{aligned}$$

where π^* is the maximized value of farm profits using a shadow wage rate, w^*, to value family labour inputs, and E is true non-labour endowment income.

The specific derivation of NLY used in the labour supply regressions is the approximation based on Y_N given by the third line of Eq. 5.3. Y_N represents the actual (not the optimized) value of the total returns to farm cultivation net of all purchased inputs, but including the implicit return to family labour applied to own farm work. The approximation for household level farm profit based on Y_N is used because Y_N is observed while Y^*_N has to be computed from the first-order conditions for a specific functional form.[23]

Since Y_N and E can be computed from the income categories reported in MPHBS, F is directly reported in the farm management module, and w^* is derived in the manner described in Section 3.5, the NLY variable based on observed data is easily computed. The optimized value of NLY would in addition be based on the estimated parameters of the production function from which the optimum levels of F and H, as well as other variable inputs, have to be computed. For methodological reasons -

[23] It may appear that the observed Y_N will consistently under-estimate the optimal Y^*_N since the latter is an extremum value and therefore Y_N on average would equal Y_N^* only if all households achieved their profit maximizing optimum input levels. However, in the presence of an additive error term in the production function, as specified in Eq. 5.1, there will be random errors in farm output, conditional on the actual input levels applied. Hence the observed Y_N can be greater or less than the optimal Y^*_N (which is computed without considering the error term). Therefore Y_N can on average be a consistent approximation for Y^*_N without resorting to the extreme assumption that all farm households achieve their profit maximizing input levels.

relating to the two step estimation strategy - it is desirable that the labour supply regressions be independent of all of the other estimated parameters of the production function, except for the unobservable w* variable (which would reflect the extent of labour heterogeneity between family and hired labour). Such a procedure can provide an independent verification of the labour heterogeneity results of the production function (and partially compensate for the fact that the production function and labour supply regressions are not jointly estimated to derive θ).[24]

The main income categories included in the E variable are household income from gifts and remittances, land rental payments and rent from other farm assets such as bullocks and farm machinery. In order to avoid having HNY become zero for most landless households (for whom $Y_N = 0$ by definition) the adopted definition of E also includes the imputed rental value of the family residence which is also reported as income for every household in the MPHBS data.[25]

5.4 Summary

This Chapter presented a brief overview of the setting of the Nepal *tarai* and the MPHBS data which is used in the estimation of the farm household model in Chapters VI and VII. Although the MPHBS data from other regions of Nepal was also available, the empirical work in this thesis is limited to the *tarai* region sample because of the limited usage of hired labour in other regions. The clustered sample design of the MPBHS give rises to significant variation in the independent variables, such as wage rates, to permit estimation of the farm household model with cross-sectional data from the *tarai* region only.

[24] Chapter VII provides further details on how the labour supply regressions can be used to choose between alternative model specifications that support labour heterogeneity or homogeneity, independently of the production function estimation results.

[25] The E income categories are derived from Section 3.1 of the MPHBS survey questionnaire Form *Ga* for the cash income components and section 5 of Form *Ga* for the income in kind categories which are received as gifts.

This Chapter also discussed in some detail the procedures and definitions followed in computing the main variables required for the regression analyses of the production and labour supply functions. The main features were that that the production function is specified at the level of aggregate farm output mainly because data on land quality are not available at the individual crop level. The labour supply equations are specified for individual family members and limited to specific agricultural occupations. Individual level wage rates are not reported in the MPHBS data. Instead the village level wage rates for hired labour derived from the farm management component of the survey are used as proxies for the off-farm wage rate for family labour. It is argued that in the setting of the Nepal *tarai* region there will be a close relationship between the wage rate reported as paid out by households hiring in labour in a specific village, and the wage rates received by individuals from that same village who report working on the hired labour market.

Appendix Table 5A.1
Distribution of Sample Households by District and Operated Land Strata

Operated Land Strata col. 1	Sample Districts by Region: Eastern (*Morang*) 2	Central (*Mahottari*) 3	Western (*Rupandehi*) 4	Mid-West. (*Banke*) 5	Far-West (*Kailali*) 6	Total Sample Size (N)
BIG	19	27	27	12	18	103
MEDIUM	26	31	31	16	19	123
SMALL	79	42	44	33	29	227
MARGINAL	82	107	57	23	12	281
ALL cultivators	206	207	159	84	78	734
LANDLESS	96	95	29	21	32	273
TOTAL	302	302	188	105	110	1007

Note: The actual sample of households used in the regression analyses in Chapters VI and VII differ slightly from the above due to additional adjustments explained in each specific chapter.

CHAPTER VI

PRODUCTION FUNCTION ESTIMATION RESULTS AND TESTS FOR HETEROGENEITY BETWEEN FAMILY AND HIRED LABOUR

6.1 Introduction

This Chapter presents the results of a production function based estimation procedure to test for heterogeneity between family and hired labour inputs in crop production in the *tarai* region of Nepal. The production function regression results also serve as the first step of the sequential estimation strategy outlined in Chapter IV for estimating a farm household model that allows for labour heterogeneity. The parameters of the production function that describe the specific nature of the labour heterogeneity will be used to generate the unobservable variables for the labour supply regressions in the second step (Chapter VII). Unlike the standard two-step model with "generated regressors" (where the first step is carried out only to derive the unobservable variables for the second step estimation), the estimated production function parameters presented in this Chapter are of independent interest. These parameters describe the production component of the farm household model, irrespective of the nature of the heterogeneity detected. They are used to derive the appropriate set of factor demand elasticities and the elasticities of input substitution that reflect the production side decision making of the farm household.

The test for labour heterogeneity is embedded in a production function structure that uses alternative ways of aggregating family and hired labour into a composite labour input. Consistent aggregation of family and hired labour, however, pre-supposes that the underlying production function is separable in the labour inputs (Berndt and Christensen 1974). Therefore it is necessary first to test for the separability of the two labour inputs in the aggregate farm production function. If the labour inputs are indeed separable, it allows a direct way of testing for labour heterogeneity together with the estimation of the complete set of parameters of the production function. In addition, as discussed in Chapter III, the separable input structure is advantageous for generating the unobservable shadow wage rates for family labour needed in the second step labour supply equations.

The statistical inference in this chapter is based on the primal estimation of a farm-level production function. The problem with direct estimation of a production function when the variable inputs could be endogenous is well known (Griliches 1984). A conventional resolution of this problem is to interpret the production function relationship in the light of the Zellner, Kmenta and Dreze (1966) framework of expected profit maximization by the farm-households, which makes the error term independent of the inputs. We adopt such a framework because the alternative of using a dual estimating procedure (based on a cost or profit function) is not feasible in this context because the shadow wage rate for family labour under a general form of labour heterogeneity is unobservable.[1]

The empirical results presented in this Chapter are based on two different sub-sets of the MPBHS sample for the five *tarai* region districts. One data set consists only of those households that utilize both family and hired labour in crop production (Sample I). The second data set consists of all farm households that report any use of family labour (Sample II). Sample I is obviously a sub-set of Sample II. These two data sets are described in Section 6.2, together with a summary table of the main variables of interest for the production function regressions.

The empirical results based on Sample I are given in Section 6.3, using a translog specification for an aggregate production function with family and hired labour as two distinct inputs. The main results are on the tests for the separability of family and hired labour. The values of the various elasticities of substitution derived from the estimated parameters are also reported for this specification. After establishing that the production function is indeed separable in these two types of labour, in Section 6.4 alternative functional forms for aggregating the two labour categories into a single composite input are tested. These tests are based on the data from Sample II that also includes households where no hired labour is used. Section 6.5 presents the full estimation results for the preferred labour aggregation functional form. Section 6.6 tests for the robustness of the labour heterogeneity result through

[1] See also the spirited defence of the primal approach to production function estimation offered in Mundlak (1996). He argues that estimating dual specifications with prices assumed exogenous at the firm level does not utilize all the available information and the loss in efficiency can be sizeable.

sensitivity analyses. The full set of the elasticities of substitution and factor demand elasticities based on the preferred composite labour function, estimated over Sample II, are derived and discussed in Section 6.7. The last section provides a summary.

6.2 Data Summary

General issues related to the structure of the MPBHS data and variable definitions adopted in this study were discussed in Chapter V. The main point relevant to this Chapter was the decision to specify an aggregate production function and to add up male and female labour inputs assuming perfect substitutability with unequal productivity.[2] This section gives further information about the actual sample of households used for the production function estimates in the different parts of Chapter VI that follow, and a summary description of the main variables of interest.

The coverage of the farm management component of the MPHBS data set is limited to households that operate some land for crop production. Ignoring the purely landless households in the sample of the 5 selected *tarai* districts, farm management data is available for 734 cultivator households. (Appendix Table 5A.1 in Chapter V gives the distribution of these households by the operated land size strata).

As noted in Chapter V, the farm management data set was processed independently of the rest of the MPHBS. Sample household identification codes were used to link the farm management data set with the data tapes received from Nepal Rastra Bank for the other parts of the survey. Due to coding errors a few households in the farm management data set could not be linked in this manner, and they have been discarded from this data set. In addition, the farm management records for several other households had to be discarded due to missing data or because the conversion factors for the local units of measurements recorded in the farm management

[2] The dependant variable in the aggregate production function is a real index of composite crop output created by deflating the nominal value by a village-specific producer price index. The sub-totals for family and hired labour add up the male and female labour days, after converting female work-days into equivalent male days, using the observed female to male wage ratio. For instance, total hired labour days = hired male days + $wr \cdot$ hired female days, where wr is the ratio of the wage rate of the female to male hired labour wage rate reported for a particular sample village. The total family labour input is computed similarly, using the wr observed for hired labour wage rates.

questionnaire were not noted in the MPHBS codebook. The actual usable sample size for the farm management data created at IIDS, after data cleaning and verifying household identification codes, was 713 households.

For the estimation work of this thesis, a final adjustment to the sample was made by discarding those households whose recorded land holding was less than 0.1 hectares or those households which did not report any use of own-family labour as inputs in crop production.[3] These adjustments give a final sample of 679 households, which is Sample II - the data set used in the main empirical work of Sections 6.4 to 6.7.[4]

For the tests of separability of family and hired labour reported in Section 6.3, the sample is limited to those households that utilize both family and hired labour. In order to test for separability it is necessary to restrict the sample to cases where both types of labour inputs are used. The number of sample households where both family and hired labour are used is 279. This constitutes Sample I, and it is obviously a subset of the 679 households in Sample II.

A summary description for the main variables used in the production function regressions in these two data subsets are given in Tables 6.1a and 6.1b. As is to be expected, Sample I, which is limited to households which use both family and hired labour in crop production, contains a higher share of big farm households. The average operated farm size in Sample I is 4.11 hectares compared to 2.67 hectares in Sample II. In Sample I, 54% of the households are from the medium and large farm size strata (as defined in the MPHBS sample design) while the corresponding

---

[3] The first of these adjustments in effect re-classifies marginal farm households with an operated land size of less than 0.1 hectares as landless households. The MPHBS strata limits are arbitrary since there is no clear line distinguishing a completely landless household and an almost landless one. It is appropriate to exclude from the production function estimation those households with very small plots of land that could be just extensions of the homestead plot. Similarly, the farm households that do not report any family labour input in crop production are also unusual. Since some of the composite labour functions used in Section 6.4 rely on a ratio format of the two types of labour, and since most households use some family labour, the cases without any family labour input are dropped from the sample. This ensures that the ratio of hired to family labour is defined for all included cases. The same sample is used for all other functional specification in order to have a consistent comparison of the results. The adjustments to the sample size due to these two arbitrary cut-offs are minor. Only 34 additional households are excluded from Sample II for these reasons.

[4] Sample selection bias that may be created by these adjustments are ignored. Nonlinear methods of estimation, coupled with the two step error correction, already complicate the estimation process.

proportion in Sample II is only 33%.[5] These percentages, however, show that a substantial proportion of the bigger farm households is completely reliant on family labour. On the other hand, almost half of the households reporting hiring labour for crop production come from the small and marginal farm size categories. Because family size and labour supply behaviour of individual family members differ, neither the incidence of hired labour use nor the ratio of hired labour in total labour input are related to operated farm size in a simple proportional manner.

Another important difference between Sample I and II is the variation in the gender composition within the family and hired labour categories. The average number of work-days of family labour is substantially larger in Sample I because of the larger farm size endowments, but the extra application of family labour is biased towards male family members. The average work-days of female family labour is almost the same in Sample I and II. Hiring a relatively higher percentage of female hired labourers makes up for the lesser application of female family labour on bigger farms. Hence the estimation results on labour heterogeneity based on Sample I can be subject to a more severe confounding influence of differences in the gender composition of the family and hired labour work days.[6]

The average wage rates for male and female hired labour are also slightly lower in Sample I than in Sample II. Although these are un-weighted average wage rates that do not fully reflect the clustered sample design of the MPHBS, it is reasonable that a lower wage rate be associated with a sample where there is a higher incidence of hired labour use. In terms of the ratio of the female to male wage rates, the two samples are very similar. This ratio (*wr)* ranges from 0.64 to 1, with a mean of about 0.85 in both samples, implying that female hired labour is, on average, 15% less productive than male workers (assuming perfect substitutability).

[5] The data on the proportion of households in the big farm category reported in Table 6.1 is based on re-classifying the large and medium farm size strata of the MPBHS sample design as big farms. Similarly, the smaller other two strata have also been collapsed into a single small farm category for the estimation work of this Chapter and Chapter VII.

[6] The share of female labour in total labour input (from both family and hired sources) is almost the same in Sample I and Sample II - about one third. The respective proportions in the family and hired categories are different. In Sample I about 28% of total family labour is female work-days but the corresponding proportion is 39% for hired labour. In Sample II the differences in gender composition are smaller - about 34% of family labour is female and about 39% of hired labour is female.

Table 6.1a: Farm Management Data Summary for Sample I*

Variable	Mean	Std Dev	Min.	Max.
	(N = 279)			
Composite farm output (kg.)	8112	7037	194	50986
Composite output price (Rs./Kg.)	2.79	0.39	2.42	4.64
Labour input variables:				
Exchange Labour work days - female	0.5	3.5	0	44
Exchange Labour work days - male	1.0	5.1	0	50
Exchange Labour days - total	1.5	8.0	0	94
Family Labour work days - female	92.7	116.2	25	858
Family Labour work days - male	233.5	208.4	32	1530
*Total Family labour work days - female	93.2	116.3	25	858
*Total Family Labour work days - male	234.7	208.2	32	1530
Hired Labour work days - female	119.0	156.9	8	960
Hired Labour work days - male	184.5	228.0	13	1295
Percentage of households using hired labour	100			
% of households using female hired labour	88			
% of households using male hired labour	92			
Share of hired labour in total labour input	0.42			
Total Female labour work days (all sources)	212.3	187.7	37	994
Total Male labour work days (all sources)	419.0	339.2	52	2345
Total Labour input (both genders) -unadjusted	631.3	485.5	95	2988
Total Labour input (both genders) - adjusted**	589.6	456.2	83	2846
Share of labour in value of production	0.28	0.09	0.07	0.63
Nominal female hired labour wage rate (Rs./day)	8.94	1.32	6	15
Nominal male hired labour wage rate (")	9.61	1.79	6.43	15
Ratio of female to male wage rate (*wr*)	0.86	0.07	0.64	1
Land input variables: (in hectares)				
Total farm cultivated area	4.11	4.08	0.1	37.4
Irrigated paddy land area	1.9	2.55	0	17.0
Unirrigated paddy land area	1.73	2.46	0	13.6
Total upland farm area	0.48	2.23	0	32.6
Gross Area harvested	6.1	5.15	0.17	37.5
Cropping Intensity Ratio	1.59	0.44	0.64	3.02
Total bullock input days	123.5	110.2	16	980
Total real material inputs (Rs.)	572.0	565.9	82	3436
Proportion of households in Big Farm category	0.54	0.5	0	1
Years of Schooling of Household Head	1.97	2.02	0	8
Distribution of Households by labour market exposure:				
percentage that are autarchic	0			
percentage with off-farm labour supply	8			

Table 6.1b: Farm Management Data Summary for Sample II*

Variable	Mean	Std Dev	Min.	Max.
	(N = 679)			
Composite farm output (kg.)	5278	5738	104	50986
Composite output price (Rs./Kg.)	2.83	0.42	2.42	4.64
Labour input variables:				
Exchange Labour work days - female	1.0	6.2	0	92
Exchange Labour work days - male	1.5	7.6	0	121
Exchange Labour days - total	2.5	13.1	0	213
Family Labour work days - female	89.9	97.7	8	858
Family Labour work days - male	185.1	172.7	11	1530
*Total Family labour work days - female	90.9	98.1	8	858
*Total Family Labour work days - male	186.6	172.8	11	1530
Hired Labour work days - female	52.9	120.0	0	960
Hired Labour work days - male	82.6	176.9	0	1295
Percentage of households using hired labour	44			
% of households using female hired labour	41			
% of households using male hired labour	43			
Share of hired labour in total labour input	0.2			
Total Female labour work days (all sources)	143.8	151.6	8	994
Total Male labour work days (all sources)	269.1	278.7	11	2345
Total Labour input (both genders) -unadjusted	412.9	402.8	21	2988
Total Labour input (both genders) - adjusted**	380.7	378.0	19	2846
Share of labour in value of production	0.29	0.1	0.07	0.82
Nominal female hired labour wage rate (Rs./day)	9.21	1.28	6	15
Nominal male hired labour wage rate (")	9.86	1.86	6	16.5
Ratio of female to male wage rate (*wr*)	0.85	0.07	0.64	1
Land input variables : (in hectares)				
Total farm cultivated area	2.67	3.21	0.1	37.4
Irrigated paddy land area	1.27	2.05	0	17.0
Unirrigated paddy land area	1.05	1.85	0	13.6
Total upland farm area	0.35	1.54	0	32.6
Gross Area harvested	4.11	4.31	0.1	37.5
Cropping Intensity Ratio	1.64	0.43	0.86	3.05
Total bullock input days	83.0	87.6	6	980
Total real material inputs (Rs.)	347.1	445.2	40	3436
Proportion of households in Big Farm category	0.31	0.46	0	1
Years of Schooling of Household Head	1.33	1.75	0	8
Distribution of Households by labour market exposure:				
percentage that are autarchic	23			
percentage with off-farm labour supply	33			

6.3 Estimation and Inference for Households Using Family and Hired labour

6.3.1 Input separability

For a production process utilizing *n* inputs, separability of inputs X_1 and X_2 from the other inputs implies that the marginal rates of substitution between X_1 and X_2 are independent of the levels of the other *n*-2 factors. That is, in the production function given by Eq. 6.1

$$(6.1) \qquad Q = f(X_1, X_2, ...X_n)$$

inputs X_i and X_j are separable from X_k if

$$(6.2) \qquad \frac{\partial \left(\dfrac{\partial f(x)/\partial x_i}{\partial f(x)/\partial x_j}\right)}{\partial x_k} = 0$$

since $\dfrac{\partial f(x)/\partial x_i}{\partial f(x)/\partial x_j}$ = the marginal rate of substitution between input *i* and *j*.

Two alternative equivalent representations of (6.2) are

$$(6.3) \qquad \frac{\partial \ln(\partial f(x)/\partial x_i)}{\partial \ln x_k} = \frac{\partial \ln(\partial f(x)/\partial x_j)}{\partial \ln x_k}$$

i.e., the elasticity of the marginal product of x_i with respect to x_k is equal to the elasticity of the marginal product of x_j with x_k (Chambers 1989: 43);

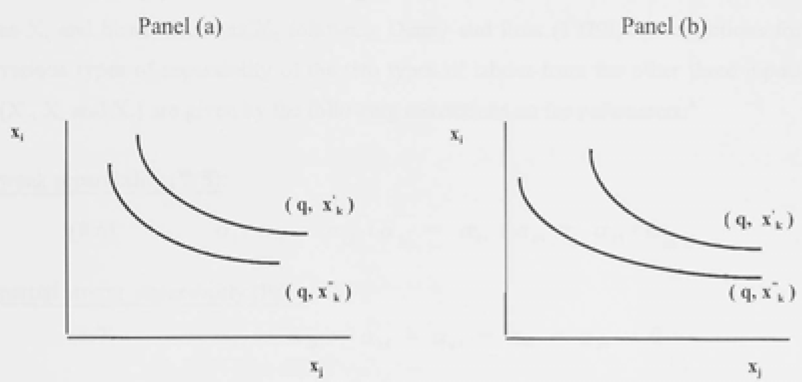
and

$$(6.4) \qquad \sigma_{ik} = \sigma_{jk}$$

where σ_{ik} is the Allen partial elasticity of substitution (AES) between input *i* and *k*, and similarly σ_{jk} is the AES between input *j* and *k* (Berndt and Christensen 1973a).

The intuitive meaning of separability implied by conditions (6.3) and (6.4) is more clearly demonstrated by considering what happens to the isoquant in i , j space when the level of input *k* is changed. If the slope of the isoquant is not affected by the level of x_k, then inputs *i* and *j* are separable from input *k*. This situation is shown in Figure 6.1(a). Changes in the application of input *k* lead to a parallel shift in the isoquants in i , j space. If there is a rotation in the isoquant when levels of input *k* are changed, as in Figure 6.1(b), then clearly the marginal rate of substitution between inputs *i* and *j* are affected by the levels of input *k*. Hence x_i and x_j are not separable from x_k.

Figure 6.1 Input Separability: Diagrammatic Representation



For the subset of the farm households that utilize some of both family and hired labour for on-farm crop production, the separability of family and hired labour from the other inputs can be tested through parametric restrictions on the production function estimated with family and hired labour as separate inputs. The translog functional form is used for this purpose as an approximation to the true underlying production function. It allows for a flexible characterization of the relationship between family and hired labour.

The translog equation to be estimated is:

(6.5) $\ln Q = \alpha_0 + \sum_{i=1}^{5} \alpha_i \ln X_i + 1/2 \sum_{i=1}^{5}\sum_{j=1}^{5} \alpha_{ij} \ln X_i \ln X_j + \sum_k \delta_k Z_k$

where family labour is indexed as input 1 (X_1), hired labour as input 2 (X_2), both measured in days; X_3 is gross harvested area, X_4 bullock days and X_5 material inputs. The elements of Z are additional explanatory variables that are not interacted with the elements of X. These include variables which control for the quality of land, such as the ratio of irrigated land and the ratio of paddy land in total holding size, and other variables such as the educational level of the household head. These are not inter-acted with the five main inputs to keep the number of estimated parameters manageable. Several dummy variables are also included in Z. The main ones are the dummies for the sample village clusters and a dummy for farm size.[7]

For the translog production function given by Eq. 6.5, with family labour specified as X_1 and hired labour as X_2, following Denny and Fuss (1979), the conditions for various types of separability of the two types of labour from the other three inputs (X_3, X_4 and X_5) are given by the following restrictions on the parameters:[8]

weak separability (WS):

(6.6) $\alpha_1 / \alpha_2 = \alpha_{13} / \alpha_{23} = \alpha_{14} / \alpha_{24} = \alpha_{15} / \alpha_{25}$

partial strong separability (PSS):

(6.7) $\alpha_{13} = \alpha_{23} = \alpha_{14} = \alpha_{24} = \alpha_{15} = \alpha_{25} = 0$

complete strong separability (CSS) (i.e. Cobb-Douglas functional form):

(6.8) $\alpha_{ij} = 0$ for all i, j (i = 1 to 5 ; j = 1 to 5)

[7] The four land size strata of the MPHBS are collapsed into a single farm size dummy variable, re-classifying the large and medium farm size strata of the MPBHS as big farms and the small and marginal holdings as small farms. Consequently, the dividing line between small and big farms is 2.73 hectares. Similarly, the 32 sample village clusters in the five districts were represented by dummy variables to capture underlying land quality differences through fixed effect intercepts.

[8] These parametric restrictions are based on the treatment of the translog flexible form as a second order approximation to an unknown arbitrary production function. Denny and Fuss (1977) call this an "approximate test " for separability. When the translog is interpreted as an exact production relationship the separability tests involve additional restrictions on the parameters, as in Berndt and Christensen (1973b). Denny and Fuss (1977:405) discuss the differences between the approximate and exact tests for separability and suggest the use of the former because the latter involves a joint test for separability and specific functional forms, and hence are unduly restrictive.

From the above restrictions it is clear that strong separability implies weak separability but not *vice versa*.[9] Finally, the complete strong separability restriction when imposed on a translog equation is not a restriction about separability alone only but also of a specific functional form.[10]

6.3.2 Test for Separability of Family and Hired Labour

The results of the statistical tests for the various forms of separability restriction are given in Table 6.2. These results are based on an ordinary least squares (OLS) estimation of the translog production function of Eq. 6.5. Two different test procedures are followed. In Test A, Eq. 6.5 is estimated without any additional restrictions (apart from the conventional one of symmetry between α_{ij} and α_{ji}). The restrictions implied by the various types of separability outlined above are then tested in separate independent tests, using the Wald Chi-square test statistic based on the unrestricted parameters.[11]

Test B follows a nested testing sequence where the subsequent tests are based on imposing the prior null hypotheses that have not been rejected. At the first level is a test for constant returns to scale (CRTS) of the five-input translog (TL) production function specified in Eq. 6.5. This restriction is not rejected. At the second level, Eq. 6.5 is re-estimated by imposing the CRTS restrictions and the test for weak separability is based on the parameters of the CRTS restricted TL equation. If weak

---

[9] The restrictions of Equation 6.6 can be re-written as

$$(6.6a) \qquad \alpha_1\alpha_{23} = \alpha_2\alpha_{13}\,;\ \ \alpha_1\alpha_{24} = \alpha_2\alpha_{14}\,;\ \ \alpha_1\alpha_{25} = \alpha_2\alpha_{15}$$

which are always satisfied when Eq. (6.7) holds. Hence partial strong separability is a sufficient but not necessary condition for weak approximate separability

[10] Another term for complete strong separability is factor wise separability (Chambers 1988:46-47). The weak and partial strong separability conditions are restrictions about the equality of the Allen elasticity of substitution between inputs in different partitions of the input list (Berndt and Christensen 1973a). The Cobb-Douglas form not only imposes equality of the AES between all input pairs but also restricts the value of this common AES to 1.

[11] The Wald test statistic is used for the tests on separability even though the equations are estimated by OLS because the tests in some instances involve non-linear restrictions on the parameters.

separability (WS) is not rejected at the second level, Eq. 6.5 is re-estimated with both the CRTS and WS restrictions imposed; and then, at the third level, the restrictions for partial strong separability are tested. If this is also not rejected, Eq. 6.5 is re-estimated with CRTS and PSS restrictions imposed, and a fourth level test for complete strong separability is implemented. The test sequence stops whenever a prior higher level null is rejected.

For nested tests the appropriate significance level of a specific test at the second and further levels depend on the significance levels of all the prior hypotheses in the sequence which were not rejected. Let H^i_0 be the i th null hypothesis in a particular test sequence to be with a assigned significance level of δ_i , conditional on the sequential non-rejection of the i -1 prior hypotheses with assigned significance levels $\delta_1, \delta_2 \ldots \delta_{i-1}$. Then the appropriate significance level for testing H^i_0 is given by (Denny and Fuss 1977)

$$1 - \prod_{j=1}^{i} (1 - \delta_j) \tag{6.8}$$

which is $= 1 - (1 - \delta)^i$, if a common significance level (δ) is assigned to each level.

A significance level of 0.025 is chosen for each level in order to get reasonable significance levels in the second and third levels. The overall test significance level is 0.040 at the second level, 0.073 at the third and 0.096 at the fourth level.

The test results reported in Table 6.2 are based on ordinary least squares (OLS) estimation with White's correction for heteroskedasticity consistent standard errors.[12] In the independent series of tests (Test A) only the WS null is not rejected. The other forms of more restrictive separability of hired and family labour are clearly rejected at the 5% significance level. The WS restriction (which has a Wald test statistic with a p-value of 0.45) is readily accepted by the data. The separate test for constant

[12] This is computed with the HETCOV option in the Shazam regression program for OLS. See Shazam (1992: 79).

Table 6.2 Testing for Separability of the Labour Inputs

Test A: Independent separate tests based on Base Equation of Unrestricted Translog

test number	test for:	signif. level	Wald test statistic	DF	p-value	Inference: Reject Null ?
i.	WS	.05	2.51	(3)	0.473	No
ii.	PSS	.05	17.01	(6)	0.009	Yes
iii.	CSS	.05	80.17	(15)	0.000	Yes
iv.	CRTS	.05	9.26	(6)	0.159	No

Test B: Sequential tests based on nested prior restrictions

Test level	Test for	Significance Levels		Wald test Statistic	DF	p-value	Reject Null?
		Individ. test	Nested test				
1.	CRTS	.025	.025	9.26	(6)	0.159	No
2.	given1, WS	.025	0.05	0.35	(3)	0.950	No
3a.	given 2, PSS	.025	0.073	9.58	(3)	0.002	Yes
3b.[a]	given 2, CSS	.025	0.073	59.6	(7)	0.000	Yes

[a] Note: The fourth level nested test sequence for CSS is not reached because PSS is rejected in the third step. Test level 3b is an alternative third level direct test for CSS based on the restrictions not rejected at levels 1 and 2 only (CRTS and WS).

CRTS = constant returns to scale
WS = weak separability
PSS = partial strong separability
CSS = complete strong separability (equivalent to Cobb-Douglas)

DF = degrees of freedom for the computed Wald Chi-square test statistic
p-value = probability value of the computed Wald test statistic.
If the p-value is larger than the significance level used for the test of the null hypothesis, the null is **not** rejected.

returns to scale, which is independent of the separability restrictions, also shows that the data does not reject the null of constant returns to scale.[13]

In Test B the test sequence begins with the CRTS restrictions since constant returns to scale is a desirable property to have in production function estimation.[14] Since test number (iv) in the Test series A did not reject the null of CRTS, this condition is imposed at the first level in the Test series B. (Test A(iv) and B(1) are identical). After imposing CRTS, the second level B test is for WS using the CRTS restricted TL parameters. Again WS is not rejected. At the third level, the estimated TL equation imposes *a priori* both CRTS and WS, and the test is for the additional restrictions of PSS. The results in Table 6.2 show this restriction is barely rejected (test level 3a), at the nested significance level of 7.3 %. The nested testing sequence properly terminates at this level. For completeness, an alternative third level test is also reported for complete strong separability - the Cobb Douglas form - given prior restrictions of CRTS and WS. The Cobb Douglas functional form is strongly rejected, with a Wald test statistic of 62.8 with 7 degrees of freedom (test 3b). The result of the nested test sequence is that the most specific restriction accepted by the data are constant returns to scale and weak separability of family and hired labour. [15]

In summary, both Tests A and B show that the family and hired labour are only weakly separable from the other three main inputs (land, bullock days and material inputs). A more restrictive form of partial strong labour separability is rejected. Also, the most restrictive form - complete strong separability, which implies a Cobb-Douglas functional form - is strongly rejected.

---

[13] The CRTS restrictions in Eq. 6.5 are imposed on the X variables only. The parametric restriction implied by CRTS in the translog equation are $\sum_{i=1}^{5} \alpha_i = 1$; $\sum_{i=1}^{5}\sum_{j=1}^{5} \alpha_{ij} = 0$ for each i = 1 to 5.

[14] The restricted estimates have a smaller variance-covariance matrix than the unrestricted estimates; so if the restrictions are accepted, it is desirable to use the restricted specification (Greene 1993: 205).

[15] The non-rejection of the null of weak separability of the two types of labour in Test B of Table 6.2 is not conditional on the prior restriction of constant returns to scale. In Test A, WS is accepted even without imposing CRTS. In Test B if WS is imposed at the first level and CRTS tested in the second level, the same overall test result occurs. The only difference is in the levels of significance of the tests at individual levels.

6.3.3 Parameter estimates

The full regression results for the CRTS and WS restricted translog production function for aggregate farm output with family and hired labour as separate inputs are given in Table 6.3.[16] Three alternative sets of parameter estimates for the TL equation are presented. The specification in Model 1 imposes only the CRTS and WS restrictions discussed above. A potential problem with the parameter estimates of a translog equation, as with other flexible functional forms, is that the theoretical curvature conditions on the concavity of production functions could be violated (Diewert and Wales 1987). Since the parameter estimates will be used to derive the elasticities of substitution and factor demands which are meaningless if the estimated production function is not concave, it is necessary first to verify that the translog specification of Model 1 satisfies the concavity property at least at the mean of the sample data. Unfortunately it does not.[17] Hence the CRTS -WS restricted specification of Model 1 is re-estimated again with additional restrictions imposed to satisfy local and global concavity conditions.[18] The results are reported in Table 6.4 as Models 2 and 3 respectively.

---

[16] The Translog regression results reported in Table 6.3, as well as in all other subsequent tables in this chapter are incomplete results. The coefficients for other variables included in these regressions are not shown. The main set of such variables was the 31 intercept dummy variables for the sample villages *(panchayat)* . The combined set of these dummy variables in the regressions results are very significant, indicating unobserved regional heterogeneity.

[17] Concavity of the TL function at specific positions can be tested by verifying the negative-semi-definiteness of the Hessian matrix of second derivatives (Berndt and Christensen 1973b).

[18] Local concavity imposes concavity at a point of approximation of the TL function - such as the geometric mean of the sample data. Global concavity requires the function to be concave at every point and this often reduces the flexibility of the translog specification (Diewert and Wales 1987). The procedures for estimating the TL equation with global concavity restrictions follow the Cholesky decomposition approach of Jorgenson and Fraumeni (1981), and Ryan and Wales (1998) for local concavity.

Table 6.3 Translog Function Parameters with Family and Hired Labour as Independent Inputs

Dependent Variable	: Log composite farm output quantity
Sample subset	: Sample I (households using family and hired labour) (N =279)
Prior Restrictions	: Constant returns to scale and weak labour separability
Estimation Method	: Non-Linear least squares

		Model 1		**Model 2**		**Model 3**	
additional restrictions:		none		local concavity		global concavity	
Variable		*Coeffic.*	*t- ratio*	*Coeffic.*	*t- ratio*	*Coeffic.*	*t- ratio*
Family lab. (F)	$\alpha 1$	0.223	*7.28*	0.227	*8.10*	0.043	*2.95*
Hired Lab (H)	$\alpha 2$	0.141	*5.62*	0.142	*5.90*	0.033	*1.56*
Land (A)	$\alpha 3$	0.447	*6.49*	0.430	*7.30*	0.765	*15.50*
Bullocks (B)	$\alpha 4$	0.104	*3.76*	0.104	*3.73*	0.092	*3.45*
Material inputs (M)	$\alpha 5$	0.083	*2.34*	0.095	*3.47*	0.067	*2.10*
Educ of Head		0.012	*1.87*	0.013	*1.86*	0.094	*1.31*
Big Farm dummy		0.041	*1.39*	0.036	*1.23*	-0.020	*0.65*
Second order coefficients (x 10)							
FxF	$\alpha 11$	0.437	*3.87*	0.442	*3.93*	-0.0002	*0.03*
HxH	$\alpha 22$	0.581	*6.45*	0.590	*6.81*	-0.0069	*0.08*
FxH	$\alpha 12$	-0.821	*4.10*	-0.838	*3.80*	0.001	*0.08*
AxA	$\alpha 33$	-3.578	*2.39*	-3.690	*2.49*	-0.081	*0.11*
BxB	$\alpha 44$	0.508	*1.14*	0.448	*1.25*	-0.016	*0.11*
MxM	$\alpha 55$	1.114	*1.74*	0.657	*2.10*	-0.053	*0.13*
FxA	$\alpha 13$	1.387	*3.08*	1.420	*3.15*	0.004	*0.05*
HxA	$\alpha 23$	0.870	*2.81*	0.890	*2.86*	-0.024	*0.02*
FxB	$\alpha 14$	-0.640	*2.40*	-0.659	*2.55*	-0.002	*0.06*
HxB	$\alpha 24$	-0.402	*2.36*	-0.410	*2.50*	0.011	*0.12*
FxM	$\alpha 15$	-0.364	*1.45*	-0.370	*1.54*	-0.003	*0.06*
HxM	$\alpha 25$	-0.229	*1.47*	-0.232	*2.47*	0.019	*0.11*
AxB	$\alpha 34$	1.188	*1.65*	1.030	*1.52*	0.036	*0.11*
AxM	$\alpha 35$	0.133	*0.20*	0.350	*0.69*	0.065	*0.12*
BxM	$\alpha 45$	-0.654	*1.62*	-0.400	*1.67*	-0.029	*0.13*
R-Sq. between observed & predicted		0.955		0.955		0.944	
Residual sum of Squares		10.31		10.34		12.91	
Log Likelihood		64.26		63.78		32.85	
Test for Cobb-Douglas form :							
Wald test statistic (with 7 DF)		59.60	(reject)	86.90	(reject)	0.19	don't reject

The parametric estimates of the CRTS-WS Translog equation, with local concavity imposed at the geometric mean, are similar to the unrestricted estimates (Model 1). The log likelihood value decreases very marginally from 63.78 to 63.69 with the imposition of local concavity; and there are only minor changes in the estimates of the parameters of the inter-action (second order) variables. Local concavity restrictions change the sign of only three of the fifteen second-order coefficients, and all three are for parameters (α_{15}, α_{25}, α_{35}) which are not significantly different from zero in the unrestricted model. In both Models 1 and 2 the α_{12} parameter which represents the inter-action between family and hired labour is significantly negative. This indicates that increasing application of the other labour input reduces the marginal product of each type of labour input, everything else held constant.

The global concavity restrictions (Model 3) however are too severe. They change the model fit and parameter estimates radically. The log likelihood value for Model 3 decreases to 32.85. All of the second order coefficients (α_{ij}) become insignificant, which reduces this model to a Cobb-Douglas specification. But even as a Cobb Douglas function the estimated elasticity parameters (α_i) are very unreasonable. They imply that the combined share of family and hired labour is less than eight percent of the value of output ($\alpha_1 + \alpha_2 = 0.076$) while the share of land is more than three fourths ($\alpha_3 = 0.765$).[19] The values of these share parameters are very reasonable in Models 1 and 2. The combined share of labour is about 37% and of land 43% while the share of bullock power and material inputs are around 10% each. Model 3 is a demonstration of the empirical problems likely to arise (as intimated in Chapter III) when family and hired labour are treated as distinct inputs, if there is a high degree of substitution between them.

[19] The data underlying the parameter estimates in Table 6.3 have been scaled at the geometric mean. Therefore the α_i represent the input elasticity at the geometric mean and are estimates of the share of input *i* in the value of total output under competitive market conditions (Boisvert 1982).

6.3.4 Elasticities of substitution

The elasticity of substitution between family and hired labour can be computed from the parameter estimates of the locally concave, CRTS and WS restricted TL in Table 6.3 (Model 2). In a multi-factor setting there are many definitions of the elasticity of substitution between any two factors. The Allen partial elasticity of substitution (AES) denotes whether two factors are *p*-substitutes or *p*-complements (where *p* stands for price). A positive (negative) AES_{ij} signifies that inputs *i* and *j* are *p*-substitutes (complements), implying that the demand for factor *i* increases (decreases) when the price of factor *j* increases, holding output constant.[20]

In a general production function with *n* distinct inputs as in Eq. 6.1, the Allen partial elasticities are given by:

$$(6.4a) \qquad \sigma_{ij} = \frac{\sum_{k=1}^{n} X_k f_k}{X_i X_j} \cdot \frac{F_{ji}}{F}$$

where f_k is the partial derivative of $f(.)$ with respect to input X_k, F is the determinant of the $(n+1)$ x $(n+1)$ bordered Hessian matrix and F_{ji} is the cofactor associated with the element f_{ji} in the bordered Hessian (Chambers 1988: 33).

Equation (6.4a) shows that the computed values of the Allen partial elasticities of substitution depend in a complicated way on all of the estimated parameters of the production function which define the bordered Hessian matrix.

The Hicksian elasticity of complementarity (HEC) between inputs *i* and *j* is defined only in terms of the first and second partial derivatives of the production function with respect to inputs *i* and *j* as:[21]

---

[20] Sato and Koizumi (1973).

[21] The Hicksian elasticity of complementarity (HEC) was introduced in Hicks (1970). It is the inverse of the Hicksian elasticity of substitution defined in the *Theory of Wages* (Hicks 1964). Sato and Koizumi (1973) have shown it is more appropriate to work with the HEC in a multi-factor setting since the HEC and AES represent the duality between the production and cost functions. In a two-factor setting, the AES and HEC are identical (Hicks 1970).

(6.4b) $$\text{HEC}_{ij} = \frac{Q \cdot f_{ij}}{f_i \, f_j}$$

Following Hicks (1970), factors i and j are classified as q-complements or q-substitutes according to whether HEC_{ij} is positive or negative (where q indicates quantity). Given non-negative marginal products, the sign of HEC_{ij} follows the sign of f_{ij} - the cross partial derivative of the production function with respect to the relevant two inputs. f_{ij} measures the effect on the marginal product of input i (f_i) due to an exogenous change in the quantity of input j. Output is not held constant in the computation of HEC (as it is in the derivation of the AES). In fact the HEC measures the degree to which two factors jointly contribute to a change in output (Sato and Koizumi, 1973).

If HEC_{ij} is positive, an increase in the quantity of factor j leads to an increase in the marginal product of input i, and so more of input i is also used, if input prices are held constant. In this case the increased use of input i and j jointly contribute to increase output. When HEC is negative, the increased usage of input j leads to a reduced demand for input i, and the combined effect on output depends on the relative magnitudes.

The direct elasticity of substitution (DES) is a multi-factor generalization of Joan Robinson's original interpretation of the elasticity of substitution as a measure of the response of the optimal factor input ratio of two inputs to a change in their factor price ratio.

(6.4c) $$\text{DES}_{ij} = \frac{\partial \ln (X_i / X_j)}{\partial \ln (f_j / f_i)}$$

The DES is a short-run measure of the substitution between two inputs, holding output and all other inputs constant (Chambers 1988:33). It ranges between zero and plus infinity, and larger values signify easier substitution between the two inputs.

The values of these various elasticities of substitution between family and hired labour, computed from the parameter estimates of Model 2 in Table 6.3 at the geometric mean of the sample data, are as follows:[22]

AES between F and H		= 10.68
HEC	"	= -1.59
DES	"	= 4.62

The positive value of AES indicates that family and hired labour are price-substitutes: as the price of one type of labour increases the quantity of the other type is increased along an isoquant. The estimated value of the AES is quite high, suggesting relative ease of substitution of one type of labour with another.[23] The negative HEC between family and hired labour indicates that increased usage of one type of labour reduces the marginal product, and hence demand for, the other. This result, together with the positive AES, indicates that family and hired labour are clearly substitutes for each other in both the price and quantity sense. The computed value of the DES between family and hired labour, which is greater than 4.6, also indicates relative ease of substitution. In a similar computation of the DES between family and hired labour in Indonesia, the highest values of the DES (among different regions) was 1.68, with many being under 1 (Squires and Tabor 1994).

The complete set of the Hicksian elasticities of complementarity for all five main production inputs at the geometric mean are reported in Table 6.4. The imposition of local concavity at the geometric means insures that the matrix of the HEC_{ij} is negative semi-definite with all of the diagonal elements (own HEC's) being negative.

---

[22] These elasticities are computed by applying Equations 6.4 (a,b,c) to derive their specific form in terms of the estimated parameters of a translog production function. These derivations are fairly involved and hence not repeated here. The computation of the AES for a translog function are detailed in Berndt and Christensen (1973b), and the derivation of the HEC and DES for a translog function are given in Squires and Tabor (1994).

[23] Note that the calculation of the AES from production function parameters involves inverting the bordered Hessian matrix of second order derivatives. Hence if any of the production function parameters have large standard errors, then the resulting estimates of the AES are imprecise, with standard errors that are not readily computable (Squires and Tabor 1994).

The imposition of weak separability of the two labour inputs (F and H) from the other three inputs also means that $HEC_{Fk} = HEC_{Hk}$, where k = L, B and M.

Table 6.4 Hicksian Elasticities of Complementarity
(with Family & Hired Labour as Separate Inputs)

	Family Lab.	Hired Lab.	Land	Bullocks	Materials
Family Lab.	-2.54	-1.59	2.46	-1.77	-0.7
Hired Lab.		-1.77	2.46	-1.77	-0.7
Land			-3.32	3.28	1.85
Bullocks				-4.46	-3.37
Materials					-2.26

Source: Derived from the parameter estimates of Model 2 in Table 6.3.

The cross-HEC's reported in Table 6.4 appear quite reasonable. Operated land area is a *q*-complement (positive HEC) for all other inputs. An increase in the land area increases the marginal product of all other inputs. Labour and bullock power are *q*-substitutes - increased application of labour reduces the marginal product of bullock power, which is a reasonable result. The least intuitive result is that labour and material inputs are *q*-substitutes. This would have been reasonable if material inputs included a lot of mechanized power inputs. But as noted in Chapter III, the main elements of M are the value of seed and fertilizers.

6.3.5 Estimated Marginal Products

The regression parameters of Model 2 (Table 6.3) can also be used to derive the marginal products of the main inputs. These are presented in Table 6.5, estimated at the geometric mean of the data in Sample I. There is a strong indication of labour heterogeneity in these results, which show a substantially higher marginal product of family labour per day (Rs.15.70), compared to the marginal product of hired labour (Rs.10.61). The estimated average marginal product of hired labour is close to the average wage rate for hired labour paid out by the sample households. The difference is not statistically significant at the 5% level, given the reported standard error on the estimated marginal product of hired labour. On the other hand, the ratio of the marginal product of hired labour to family labour (θ* in the notation of Chapter III) is 0.676. This estimate is significantly less than 1, even at the 1% level of significance.[24]

The estimates of the marginal products for the other inputs are also very plausible. Note that the marginal product of land is expressed as per hectare of gross cultivated area. This is the return to an additional hectare of multiple cropping which is not the same as the return to increasing the operated land of the farm household by a hectare. The estimated marginal product for a bullock pair per day is higher than the marginal product of hired labour, which is consistent with the observed market wage rates for these two inputs.

Table 6.5 Estimated Marginal Products of Inputs

Inputs	*units*	*Marginal Product* Value	tand. error
Family Labour (F)	Rs. per day	15.7	1.93
Hired Labour (H)	Rs. per day	10.61	1.8
Land (A)	Rs. per ha.	2371	325
Bullocks (B)	Rs. per day	19.1	5.12
Mat. inputs (M)	Rs. per unit of expenditure	4.35	1.25
Memorandum items:			
Average daily wage rate for male hired labour		9.61	
Ratio of the marginal product of hired labour to family labour		0.676*	0.083
		* significantly less than 1	

Source: Marginal products computed at the geometric mean of the data using the regression parameters of Model 2 in Table 6.3

6.3.6 Alternative Estimation with a homogeneous labour aggregate

The TL equation with family and hired labour as separate inputs gives reasonable estimates for the production function parameters as well as the marginal products of

[24] The Wald statistic for the test that the ratio of marginal products is equal to one is 15.3, with 1 degree of freedom. The p-value is 0.0001.

Table 6.6 Translog Production Function Parameters with Homogeneous Labour

Dependent Variable	: Log real aggregate farm output		
Sample subset	: Sample I (households using family and hired labour) (N =279)		
Prior Restrictions	: Constant returns to scale		
Estimation Method	:Ordinary least squares with heteroskedasticity consistent errors		
		Model 4	
additional restrictions :		local concavity	
Variable		*Coefficient*	Asymp. *t- ratio*
Total labour (L)	$\alpha 1$	0.367	5.78
Land (A)	$\alpha 3$	0.449	6.26
Bullocks (B)	$\alpha 4$	0.105	3.87
Mat. inputs (M)	$\alpha 5$	0.079	2.75
Educ		0.01	1.65
Big Farm dummy		0.024	0.9
Second order coefficients (x 10)			
L x L	$\alpha 11$	-1.35	0.64
A x A	$\alpha 33$	-6.08	2.51
B x B	$\alpha 44$	0.37	0.95
M x M	$\alpha 55$	0.66	1.07
L x A	$\alpha 13$	3.95	1.97
L x B	$\alpha 14$	-1.83	2.13
L x M	$\alpha 15$	-0.77	1.00
A x B	$\alpha 34$	1.74	1.74
A x M	$\alpha 35$	0.38	0.497
B x M	$\alpha 45$	-0.27	0.73
Adj. R- SQ.			0.955
Residual sum of squares			10.22
Log Likelihood			65.4
Test for Cobb-Douglas form			
Wald test statistic (Chi-square with 6 DF)			12.27 reject
Test for common production technology **			
F test statistic (with 48 and 581 DF)			1.36 don't reject

** This is a test for common production function paramaters in the sample of farms that use both family and hired labour (Sample I) and farms that use only family labour.

the inputs, and the various elasticities of substitution. The reasonableness of these results, however, needs to be checked against alternative specification that may be preferred to the one that has F and H as distinct inputs. Table 6.6 presents the parameter estimates for a TL specification that assumes that family and hired labour are homogeneous inputs. In this specification (Model 4) there are only four main inputs which are inter-acted with each other: total labour (L, which is simply the sum of F and H), land, bullock days and material inputs. Constant returns to scale and local concavity are imposed as in Table 6.3.

A comparison of the parameter estimates between Model 2 and Model 4 reveals only minor changes in the first order coefficients (the α_i 's). The coefficient α_l on total labour in Model 4 is almost an exact sum of the corresponding coefficients on family and hired labour in Model 2. But there are bigger changes in some of the second order coefficients (α_{ij} 's). The coefficient on the total labour quadratic term (L x L) changes sign in Model 4 and there is a large change in the coefficient for (A x A).

The effect of these changes in the estimated parameters on the underlying household behaviour can be illustrated by computing the values of the Hicksian elasticities of complementarity (HEC) for Model 4. These are reported in Table 6.7 (computed at the geometric mean of Sample I). Comparing the values of the HEC's in Table 6.7 and Table 6.4 shows that the input relationships based on a model with aggregate homogeneous labour does differ considerably from the relationships based on treating family and hired labour as separate inputs.

Table 6.7 Hicksian Elasticities of Complementarity (with Aggregate Homogeneous Labour)

	Total Lab.	Land	Bullocks	Materials
Total Lab.	-2.73	3.4	-3.76	-1.67
Land		-4.24	4.68	2.08
Bullocks			-5.17	-2.31
Materials				-1.02

Source: Computed at the geometric mean of the data using the regression parameters of Model 4 in Table 6.6

Table 6.6 also reports on a test for common production function parameters in the sample of farms that use both family and hired labour (Sample I) and farms that use only family labour (Sample II). There are no *a priori* reasons to believe that the production technology available to households that use hired labour is different from households using only family labour. This proposition is tested with the homogeneous labour aggregate because with this specification the model can be estimated by ordinary least squares for which the F test for the stability of parameters in different samples is exact.[25] The total number of parameters estimated with the homogeneous labour aggregate is 47 (including 31 sample village cluster dummies). The computed F test statistic is less than the critical value at the 5% significance level. Hence, the null of a common production technology for all farm households , irrespective of whether they use hired labour or not, is not rejected.

Returning to the parameter estimates relating to Sample I households only, Model 4 in Table 6.6 and Model 2 in Table 6.3 are not nested within each other. Therefore the evidence on which specification is preferred is not straightforward. As an illustration it is interesting to note that the standard goodness of fit criteria for non-linear regressions, such as the log likelihood and the residual sum of squares, indicate that Model 4 with the homogeneous labour aggregate provides a better fit for the sample data. This is a surprising result given that Model 2 is a flexible functional form which provides for a general form of inter-relationship between family and hired labour. If family and hired labour were truly homogeneous inputs, Model 2 could have reflected such a relationship instead of the estimated wide divergence in the marginal products of the two labour inputs at the mean of the data.

It appears that production functions that treat family and hired labour as distinct inputs could be mis-specified when in fact the two types of labour are very close substitutes. This was clearly evident in the Cobb-Douglas specification with family

[25] Tests for common production function parameters based on the nested production function structure with an effective labour nest, as specified in Section 6.4 to follow, are awkward to carry out. This is so not only because of the non-linear estimation procedures for which the F test is only approximate (Greene 1993:218); but also because the parameter set between samples using hired and family labour and those using only family labour will differ in the nested production function.

and hired labour as distinct inputs - the parameter estimates in Model 3 (Table 6.3) were nonsensical. Similar problems may carry over to more general functional forms as well, given that Model 4 with a homogeneous labour aggregate performs better than a flexible specification with family and hired labour as distinct inputs. The preferred alternative is to look at ways of aggregating the different labour inputs into a composite labour aggregate without imposing homogeneity. This is the issue explored in the next section of this Chapter.

6.4 Testing for Alternative Aggregates of Family and Hired Labour

The separability results of Section 6.3.2 imply that family and hired can be consistently aggregated into a single composite labour input. Specifying the production function with a composite labour input will also mean that the parameter estimates can be based on Sample II which also includes households that do not report any hired labour use for crop production. Such households constitute almost 60% of the total available sample of 679 land cultivating farm households created from the MPBHS *tarai* sample. The production technology available to households that use hired labour is unlikely to be different from those using only family labour; and the F test reported in Table 6.6 did not reject this null hypothesis. It is then appropriate to estimate a common production technology for all households, specified on the basis of a composite labour input, irrespective of whether or not a particular household reports any use of hired labour.

An alternative specification of the translog production function which still maintains family and hired labour as distinct inputs for the entire Sample II is inappropriate. In this case a large proportion of households would have zero inputs of hired labour. *Ad hoc* procedures of re-scaling the data by converting the zero values to one (or ten) so that logarithmic values can de defined, while commonly used, are inappropriate for this data set because of the high proportion of zero values for hired labour.[26]

---

[26] Adopting such a procedure leads to substantial downward bias in the linear coefficients of hired labour (α_2 in Table 6.3) and, hence in the implied ratio of marginal productivity of family and hired labour. See Appendix 6.1 for a fuller discussion, and in particular Appendix Table 6.13.

The customary approach to creating a composite out of two or more separable inputs is to create a value-share weighted linear or log-linear (Tornquist) index. For an underlying translog production process, a Tornquist composite of separable inputs would be an exact index (Diewert 1976). This procedure is not feasible here because zero values of hired labour are common and also because the relevant "price" of family labour is not observed. Indeed the main purpose of carrying out the the production function estimation is to decipher whether and how this (shadow) price of family labour differs from the observed market wage rate.

Consequently, alternative functional forms are specified which create a composite quantity of family and hired labour that allows for zero values of hired labour. The composite labour input then becomes a single input in the aggregate production function. The specification for the composite labour aggregator function is nested into the translog production function of Eq. 6.5, which is now re-specified in terms of the four main inputs that are interacted: composite labour, land, bullocks and material inputs. The composite labour variables can be viewed as the total input of labour in effective units (as opposed to standard time inputs of days or hours) and can be termed as effective labour.

6.4.1. Effective labour functions

Five specific functional forms are chosen to create an effective labour (Le) input from the observed levels of family and hired labour - the *g* (.) function in the terminology of Chapter III. These allow for different possibilities of efficiency differences as well as for a constant or varying elasticity of substitution between family (F) and hired labour (H) in the "production" of effective labour. These functional forms, which are adapted from Deolalikar and Vijverberg (1983), are:

6.5. C1. Homogeneous labour: $Le = F + H$

6.5. C2. Linear heterogeneous: $Le = F + \theta^{*}H$ $\quad \theta > 0$

6.5 C3. CES:[27] $Le = [F^{-\rho} + \theta * H^{-\rho}]^{-1/\rho}$

$\theta > 0,\ \rho \geq -1$

6.5 C4. Generalized Linear (GL): $Le = F + \theta * H + 2 * \delta * (H * F)^{1/2}$

$\theta > 0$

6.5 C5. Ratio: $Le = (F + H) * (F / (F + H))^{\mu}$

$|\mu| \leq 1$

Once input separability is established for a subset of inputs, there are two additional technical properties of the *g* (.) aggregator function that determine whether two different production inputs are homogeneous or heterogeneous:

(i) whether *g*(.) has a linear form such that the input components are perfect substitutes for each other, implying an infinitely large AES

(ii) whether the ratio of marginal products (the marginal rate of substitution between input pairs) is equal to one.

The five functional forms specified above allow for different possibilities on points (i) and (ii) above. Forms C.1 and C.2 imply an infinite elasticity of substitution between F and H. The CES composite implies a constant elasticity of substitution. The Generalized Linear and the Ratio form imply a varying elasticity of substitution that depends on the levels of F and H and the parameters.

The CES functional form is normally specified with the restriction that $\rho \geq -1$ which is required for concavity to ensure the isoquants are well behaved. When the CES specification is a complete production function in itself (without being nested into another functional form) concavity implies that all the inputs be *p*-substitutes with positive AES. This requires that $\rho \geq -1$ because when $\rho < -1$ the AES between the

---

[27] The traditional form of the CES function is $\gamma [\delta * F^{-\rho} + (1-\delta) * H^{-\rho}]^{-1/\rho}$, where γ is the efficiency parameter and δ the distribution parameter (Arrow, et al. 1961). This can be re-specified as $\gamma' [F^{-\rho} + \theta * H^{-\rho}]^{-1/\rho}$ where $\gamma' = (\delta^{-1/\rho}) \gamma$ and $\theta = (1-\delta)/\delta$. Since γ' is just a scaling factor which gets converted to a constant when taking logs, the CES form can be written as in C.3 above. Similarly in the GL specification of C.4, the coefficient for F can also be normalized to one.

inputs are negative, implying the inputs are p-complements.[28] But when a CES aggregator function such as C.3 is nested within a higher level production function, such as the translog of Eq. 6.5, the restriction that $\rho \geq -1$ in the CES nest is not strictly required. The appropriate theoretical restriction is that the overall production function be concave. This condition depends also on the other parameters of the translog function, and hence can be satisfied even if in the CES nest $\rho < -1$.

In the empirical results given below the CES composite for family and hired labour is estimated both with and without restriction that $\rho \geq -1$. The restriction on ρ has a central bearing on the nature of labour heterogeneity that is the subject of interest. When the CES composite is estimated with the restriction $\rho \geq -1$, it forces F and H to be q-complements - i.e., an increase in the input level of F necessarily increases the marginal product of H. Hence F and H jointly contribute to increase output in the sense captured by a positive HEC.[29] But if F and H are to be q-substitutes, as is more likely, the CES form must allow for $\rho < -1$.

The Ratio composite is equivalent to Revankar's Variable Elasticity of Substitution (VES) production function, for which it can be shown that the (AES) $\sigma_{FH} = 1 + (F/H)^*(1/\mu)$ (Revankar 1971). For $\mu = 0$, this form reduces to the homogeneous labour case (C.1) with an infinite elasticity of substitution. For non-zero μ, the AES between hired and family labour is increasing or decreasing in the ratio of F to H, depending on whether μ is positive or negative.[30] That the elasticity of substitution between F and H could be related to the ratios of F and H is consistent with a particular approach to modeling labour supervisory/monitoring costs on the premise

---

[28] A multi-factor non-nested CES production function has the undesirable property that the elasticity of substitution for every pair of inputs is exactly the same (Uzawa 1963). The second order conditions for the concavity of a multi-factor production function with n inputs requires that at least (n-1) of the Allen partial elasticities of substitution must be positive (Sato and Koizumi 1973). Therefore in a multi-factor CES, all inputs must be p-substitutes with positive AES.

[29] From Equation 6.4b, the sign of HEC_{ij} depends only on the cross-derivative f_{ij}. In the CES nest of C.3, f_{FH} is always positive when $\rho \geq -1$.

[30] The restriction $|\mu| \leq 1$ is needed to maintain concavity of the ratio aggregator function.

that the hired labourer's work intensity or effort can be increased if family labour is also working alongside with hired labour (Frisvold 1994).

For the GL functional form, the AES between F and H is given by: [31]

$$(6.6) \quad \sigma_{FH} = 1 + (F + \theta H) / Le + 2 * (H*F)^{1/2} / (\delta * Le)$$

In this formulation, the elasticity of substitution depends in a more complicated way on the actual levels of F and H, and not just on their ratio. In Eq. 6.7 σ_{FH} can be negative only if $\delta < 0$ (though this is only a necessary and not sufficient condition). The concavity of the GL form can be violated when $\delta < 0$ (Diewert 1971). Again this restriction is not imposed since the overall translog production function can still be concave with $\delta < 0$.

For the set of functional forms specified above, it is clear that with appropriate parametric restrictions on ρ and δ, both the CES and the GL forms reduce to the linear heterogeneous composite (C.2); and the latter obviously reduces further to the homogeneous aggregator when $\theta = 1$. This means several combinations of the first four specifications can be tested in a nested framework. The ratio composite is unrelated to the GL, CES or linear forms; it reduces only to the homogeneous case.

6.4.2. Effective land function

In addition to the composite function for effective labour, the land input variable is also specified as a separable nest allowing for heterogeneity between different types of land and also allowing for variation in land use due to differences in the multiple cropping intensities of farms.

The effective land specification is:

[31] In a nested production function structure it is necessary to distinguish the AES between F and H in the g(.) function for the composite Le, and the AES between F and H in the main (first level) production function. Eq. 6.6 above gives the formula for the AES in the Le function of the GL specification (Deolalikar and Vijverberg 1987). The Le function has only two inputs. In a two-input production function the AES is equal to the HEC (Hicks 1970); so σ_{FH} can be computed using Eq. 6.4b for the Le specification in C.5.

(6.7) $Ae = A_c \cdot CRINT^{\kappa}$

where A_c is the physical cultivated area of the farm or the land endowment available for crop production. CRINT is the cropping intensity ratio which is equal to A/A_c, where A is the gross area harvested of different crops, which is the sum of the land area allocated to each crop grown on the farm.

This specification treats the increases in the physical area of the farm differently from increases in the harvested area due to more intensive multiple cropping. When $\kappa = 1$, Eq. 6.7 is just an identity with effective land being the same as A. If as expected $\kappa < 1$, the returns from increased multiple cropping are less than the returns from increasing the physical land area of the farm. This is consistent with decreasing returns to multiple cropping given a fixed land area.[32]

The cultivated land area available to the farm is further decomposed into three different land categories, assuming a linear form. This assumes perfect substitutability but which allows for differing productivity, as in the case of the labour inputs. The farm management data set of the MPHBS actually identifies four different types of land: irrigated paddy land (*khet*), irrigated upland (*pakho*), and un-irrigated *khet* and un-irrigated *pakho*. In the final specification for the effective land input function the distinction between irrigated and unirrigated land in the upland category did not turn out to be important and is ignored.[33]

The final specification selected for A_c is

(6.8) $A_c = A_c^{Pi} + \gamma_1 \cdot A_c^{Pu} + \gamma_2 A_c^{U}$

were P_i is irrigated paddy land, P_u is unirrigated paddy land and U is total upland. This specification means γ_1 and γ_2 are conversion factors which measure A_c in units of total irrigated paddy land. The expected relationship is< $0 < \gamma_2 < \gamma_1 < 1$ since even un-irrigated paddy land is usually more productive than upland, whether or not the latter is irrigated.

---

[32] Since the dependent variable in the production function regressions is aggregate output, allowing for $\kappa < 1$ is also an implicit control for variations in the cropping pattern that arise from differences in the level of multiple cropping, in contrast to differences in the physical land area across farms.

[33] This is a reasonable result because upland does not generally have a high water retention rate and hence unirrigated land dependent on rainfall may be just as productive as upland which has a more dependable irrigation source.

Just as in the case of the Le function for effective labour, the Ae function for effective land input is nested in the overall translog production function of Eq. 6.5, which is now defined in terms of four main inputs.

The estimating equation is

$$(6.9) \quad \ln Q = \alpha_0 + \sum_{i=1}^{4} \alpha_i \ln X_i + 1/2 \sum_{i=1}^{4}\sum_{j=1}^{4} \alpha_{ij} \ln X_i \ln X_j + \sum_{k} \delta_k Z_k$$

where X_1 is effective labour, X_2 is effective land and X_3 and X_4 are bullock power and material inputs. As before Z is the set of other variables not inter-acted with the set of X.

6.4.3. Labour heterogeneity test results

Equation 6.9 was estimated by non-linear least squares for each of the five different effective labour composites (C1 to C5), and the common effective land composite.[34] Table 6.8 presents the main results of interests, focussing on the parameters of the labour nesting function which determine the nature of labour heterogeneity, and on summary statistics for the fit of a specific model.

The main inference to be drawn from Table 6.5 is that among the set of nested models (C4 → C2 → C1 and C3 → C2 → C1), the restrictions implied by the general linear composite of model C2 are accepted. But the additional restriction that makes family and hired labour homogeneous inputs (Model C1) is clearly rejected. In comparison to the simple linear composite ($Le = F + \theta H$), the more general form of a CES composite (C3) or a GL composite (C4) do not offer any additional explanatory power to the production function regressions.

[34] In logarithmic form for the overall translog function, the specification for the effective land variable becomes $\log(Ae) = \log(A_e^{Pi} + \gamma_1 A_e^{Pu} + \gamma_2 A_e^{U}) + \kappa \log(CRINT)$.

In the CES nest with ρ unrestricted, the estimated value of ρ is -1. 03, but this is not significantly different from -1.[35] This reduces the CES function to a linear composite. Similarly in the GL specification the δ parameter is not significantly different from zero; hence, the GL also reduces to the linear nest.

At the next step, however, there is a significant difference between the linear composite model and the homogeneous model because the θ parameter is significantly less than one. The estimated value of θ using Model C2 is 0.751, with a standard error of 0.107. The Wald Chi-square test statistic for the hypothesis that $\theta = 1$ is 5.44, with a p-value of 0.03. Hence the null of $\theta = 1$ is rejected at the 5% significance level. The same inference is obtained from the Likelihood Ratio test in comparing Model C1 and C2.[36]

In the ratio format the estimate of the μ parameter is positive - which is consistent with family labour being more productive than hired labour.[37] But its estimated value of 0.042 is not significantly different from zero. Therefore the ratio format can also be rejected. Although the ratio and the linear composite (C5 & C2) are not nested, the latter is clearly preferable in terms of the log likelihood value.[38]

Hence, the preferred labour aggregator function is $Le = F + \theta H$, where θ is significantly less than one. This implies that, although family and hired labour are perfect substitutes, there is a constant efficiency difference between the marginal

---

[35] In the restricted version of the CES nest, the boundary limit on $\rho \geq -1$ is reached, So this is equivalent to estimating the CES nest with $\rho = -1$ imposed.

[36] Based on the log likelihood values reported for Models C.1 and C.2 in Table 6.8, the likelihood ratio test statistic is $-2 * (138.56 - 141.06) = 5.02$. The critical value of the Chi-square test statistic with 1 degree of freedom is 3.8. Therefore this test also shows that the sample data is not consistent with the restriction that $\theta = 1$.

[37] A positive μ implies that for a given level of aggregate labour (F + H), a higher ratio of family labour in this total increases the level of effective labour input. A positive μ also implies the AES_{FH} is positive, implying F and H are *p*-substitutes and *q*-complements.

[38] The ratio specification and the linear composite have the same number of parameters. Therefore a smaller residual variance, which leads to higher value of the log likelihood, is equivalent to that model being preferred on several model selection criteria, such as the Akaike Information Criteria, which depend only on the residual sum of squares and degrees of freedom (Shazam 1993:16).

Table 6.8 Summary Results on Tests for Alternative Composite Labour Functions in Sample II

Model	Composite Function	Parameters of interest	Estimated value	Stand. error	Tests of parametric restrictions: Restriction	Wald test statistic	p value	Inference	Summary statistics on Model fit: R sq *	Log likelihood
C4	GL	θ	0.751	0.107	$\theta = 1$	5.44	0.03	Reject	0.965	141.06
		δ	0.0075	0.087	$\delta = 0$	0.007	0.93	Do not Reject		
C3	CES (ρ restricted)									
		θ	0.780	0.114	$\theta = 1$	3.68	0.05	Reject	0.965	141.06
		ρ	-1*	-	$\rho = -1$	-	-	*Restriction imposed		
	CES (ρ unrestricted)									
		θ	0.757	0.115	$\theta = 1$	4.46		Reject	0.967	141.07
		ρ	-1.029	0.144	$\rho = -1$	0.04	0.84	Do not Reject		
C5	Ratio	μ	0.042	0.034	$\mu = 0$	1.52	0.23	Do not Reject	0.965	139.34
C2	Linear	θ	0.751	0.107	$\theta = 1$	5.42	0.02	Reject	0.967	141.06
C1	Homog-eneous	$\theta = 1$							0.964	138.56

*R-sq. between observed and predicted values.

Inference on Nested Models :

Restrictions on Model C4 and Model C3 which reduce them to Model C2 are **not rejected**
The restriction which reduces Model C2 to C1 is rejected .
The restriction which reduces Model C5 to C1 is **not rejected**.

product of family and hired labour, with family labour being more productive. It is worth noting that all of the specifications which allowed F and H to be *p*-complements in production with a negative AES (the GL with $\delta < 0$ the Ratio form with $\mu < 0$ and the unrestricted CES with $\rho < -1$) were clearly rejected.[39]

The above result that family labour is more productive per unit of time is consistent with the empirical findings of Frisvold (1994) and Deolalikar and Vijverberg (1983) for India. However, the preferred aggregator functional form is different. Deolalikar and Vijverberg (1983 and 1987) reject a linear composite and hence find evidence for imperfect substitutability between family and hired labour. Foster did not actually test alternative specifications of the Le function. He just used the ratio format and found μ was positive and significantly different from zero.[40] This also implies imperfect substitutability.

6.5 Complete results for the linear labour composite

Table 6.9 gives the complete regression results for the linear effective labour composite (Model C2 in the set of Eq. 6.5). The only prior restriction imposed is of constant returns to scale. The set of estimated parameters makes this specification concave at the geometric mean so concavity restrictions have not been imposed *a priori* (as was done in Model 2 of Table 6.3).

The set of parameters of the TL production function given in Table 6.9 appears reasonable. At the geometric mean of the data the estimated input elasticities are 0.327 for effective labour, 0.487 for effective land, 0.076 for bullock power and 0.111 for material inputs. These are reasonable estimates of the share of these inputs in the value of farm output at the mean of the data. Several of the second order

[39] In the unrestricted CES case, the estimated of ρ is less than minus one but not significantly different from minus 1. This reduces the CES to a linear composite with an infinite AES.

[40] Frisvold's estimate of μ was 0.24. At the sample mean this leads to an AES between F and H of 1.71 in the effective labour function and 0.33 in the main production function (Frisvold 1994:230-231).

Table 6.9: TL Production Function Parameters with Linear Heterogeneous Labour

Dependent Variable	: Log real composite farm output
Sample subset	: Sample II (all households using some family labour) N = 679
Prior Restrictions	: Constant returns to scale
Estimation Method	: Non-Linear least squares (heteroskedasticity corrected errors)

		Model C2 $Le = F + \theta H$		
Variable		*Coefficients*	*Asymp. std. error*	*Asymp. t- ratio*
Effective labour (Le)	$\alpha 1$	0.327	0.0439	7.45
Effective Land (Ae)	$\alpha 2$	0.487	0.0508	9.59
Bullocks (B)	$\alpha 3$	0.076	0.0204	3.73
Material inputs (M)	$\alpha 4$	0.111	0.0218	5.09
Educ. of Head		0.007	0.0051	1.37
Big Farm dummy		0.029	0.0217	1.34
Second order coefficients				
Le x Le	$\alpha 11$	-0.051	0.1048	-0.49
Ae x Ae	$\alpha 22$	-0.21	0.1465	-1.43
B x B	$\alpha 33$	0.033	0.0147	2.24
M x M	$\alpha 44$	0.067	0.0423	1.58
Le x Ae	$\alpha 12$	0.146	0.1098	1.33
Le x B	$\alpha 13$	-0.077	0.0427	-1.80
Le x M	$\alpha 14$	-0.018	0.0522	-0.34
Ae x B	$\alpha 23$	0.079	0.0472	1.67
Ae x M	$\alpha 24$	-0.015	0.0602	-0.25
B x M	$\alpha 34$	-0.034	0.027	-1.26
Labour nest parameters				
	θ	0.751	0.1068	7.03
Land nest parameters				
	γ_1	0.923	0.0461	20.02
	γ_2	0.836	0.0793	10.54
	κ	0.743	0.0653	11.38
R-Sq. between observed & predicted		0.967		
Residual sum of squares		23.71		
Log Likelihood		141.06		
Wald chi-square test statistics on restrictions				
labour homogeneity	$\theta = 1$ df =1	5.437		reject**
land homogeneity	$\gamma_1 = \gamma_2 = \kappa = 1$ df = 3	11.56		reject**
Cobb Douglas form	df = 6	12.27		reject*

Note: ** indicates the null is rejected at the 5% significance level, * at the 10% level.

coefficients are statistically insignificant, but the joint test that the model reduces to a Cobb-Douglas specification is rejected. The estimates of the parameters of the land composite function are also as expected. The coefficient on the cropping intensity component of total land use (κ) is less than one. The γ_1 and γ_2 parameters indicate that un-irrigated paddy land has about 92% of the productivity of irrigated paddy land, and upland has about 84% of the productivity of irrigated paddy land.[41] Each of the κ, γ_1 and γ_2 parameter is individually significantly less than one and the joint test for land homogeneity ($\kappa = \gamma_1 = \gamma_2 = 1$) is also rejected at the 5% significance level.

The main implication of the linearly heterogeneous labour composite is a constant difference in the marginal productivity of family and hired labour represented by θ. The implied values of the marginal products based on the regression parameters and computed at the geometric mean of the data are given in Table 6.10. For comparison the marginal product of homogeneous labour (based on the composite function C.1) is also reported.

From the estimates for the C1 homogenous labour composite, the marginal product of a homogenous labour aggregate, at the mean of the input variables, is Rs. 12.34 for an additional work-day of a male labourer. This estimated marginal product is about 25% higher than the average male wage rate in the sample data, which is Rs. 9.91 per work-day. This is a relatively large discrepancy in a setting where the normal presumption would be that the marginal product of labour would be lower

[41] These estimated relative differences in unit productivity of unirrigated paddy land and upland *vis-a-vis* irrigated paddy land appear to be smaller than expected, given the primacy of paddy cultivation in the *tarai* region of Nepal. One factor contributing to the low relative productivity differences is that the land input variable is already measured in terms of gross harvested area and not in terms of physical units of land. If irrigated land tends to be multi-cropped more often, and if the yield in the second or subsequent crop grown outside of the normal seasonal rotation is likely to low, this will depress the average productivity of irrigated land (compared to the yield of unirrigated land which is cultivated only during the normal season). Secondly, the MPHBS survey data does not contain details on the quality and reliability of irrigation facilities on the sample of farms. Since the timing of irrigated water supply is the critical issue in boosting crop yields, it may be that differences in the yield between irrigated and unirrigated land may be more pronounced in years where there is general drought due, say, to the monsoon failure. In a normal year, as in our sample, when the monsoon rainfall was adequate, the relative differences in yields could be less extreme if adequate rainfall has occurred at critical times in the growing season.

than the wage rates reported to be paid out to hired labour. Apart from a general indication of surplus labour in the *tarai* region of Nepal, the specific wage rates derived from the MPBHS data are likely to reflect peak-season hired labour wage rates than an annual average wage rate.[42]

The model with the Le $= F + \theta H$ specification helps to reconcile the observed gap between average labour productivity and wage rates. With this specification the estimate of the marginal productivity of hired labour at the mean of the data is only Rs. 9.19 per male work day. This is slightly lower than the average wage rate. This result is more in conformity with profit maximizing behaviour of the sample households. Whenever hired labour is used, the estimated marginal product on average is close to the market wage rate for hired labour.

The effect of the efficiency difference between family and hired labour is that it makes the effective return to family labour higher in own farm production as a substitute for hired labour. Family labour in such situations should act as if the effective wage rate for own-farm work is higher than the market wage rate paid to hired labour. This should induce a greater labour supply response of family members in big farm households. Whether the actual labour supply behavior of family members in households that hire in labour is consistent with the specific form of labour heterogeneity indicated in the production function estimation with $\theta = 0.75$ is taken up in Chapter VII.

[42] As described in Chapter V, the wage rates are compiled from the actual wage payments made by households that report hiring in labour for each sample village cluster. Since labour hiring in is likely to be more common in the peak labour demand seasons, these wage rates will tend to be higher than an annual average of the hired labour wage rates, when there is seasonal fluctuation in the wage rates.

Table 6.10 Estimated Marginal Products of Labour

Model	*Inputs*	*units*	*Marginal Product* Value	tand. error
Le = F + H				
	Labour	Rs. per day	12.34	1.51
Le = F + θ H				
	Family Labour (F)		12.24	1.64
	Hired Labour (H)		9.19	2.12
Memo item :				
	Average wage rate for male hired labour		9.86	1.86 *

Source : Marginal products computed at the geometric mean of the Sample II data using regression parameters of Models C.1 and C.2
* Standard deviation of male wage rate in sample

6.6 Sensitivity Analysis

Given the critical role that θ plays in the characterization of labour heterogeneity on the production side of the farm household model, and in the subsequent specification for the labour supply equation, it is important to establish the robustness of the result that θ is significantly less than one. Table 6.11 presents a sensitivity analysis of the estimates of θ from alternative restrictions placed on the translog production function with the linear heterogeneous nest for effective labour. In all specifications reported in Table 6.11, the null that $\theta = 1$ is rejected at the 5% significance level.

In particular it is heartening that the estimate of θ appears to be quite insensitive to the conversion factor between male and equivalent female work-days. In the main production function estimation results reported in Sections 6.3 to 6.5 of this Chapter, this conversion factor has been the ratio of the female to male wage rate in each sample village (*wr*). The mean value for *wr* was 0.85 with a range from 0.64 to 1. In the results reported under item 7 of Table 6.11, the TL equation has been re-run by replacing the actual reported value of *wr* in each sample village with an exogenously specified value. Whether this exogenously specified common value of *wr* is 0.6 or 1, the estimate of θ changes only marginally. More importantly the result that the estimated θ is significantly differ from 1 is not altered by a wide variation in *wr*. Therefore, as long as a linear specification implying perfect substitutability between male and female labour is imposed *a priori*, the specific rate at which female labour work-days is converted to equivalent male work days does not matter. It does not affect the main result that family and hired labour for both sexes are perfect substitutes with a constant higher productivity for family labour.

Table 6.11 Estimates of the θ parameter of the Linear Effective Labour Nest under alternative functional forms and prior restrictions

	(Le = F + θ*H)		θ	stand. error	Testing Null θ = 1 Wald Statistic	Inference (5% signif.)	Log likelihood
1	Cobb-Douglas	unrestricted	0.718	0.116	5.91	reject	127.7
2	Cobb-Douglas	with CRTS	0.706	0.119	6.1	reject	126.3
3	Translog	unrestricted	0.708	0.109	7.18	reject	143.9
4	Translog	with CRTS	0.751	0.106	5.52	reject	141.1
5	Translog	with Local concavity	0.708	0.109	7.18	reject	143.8
6	Translog	CRTS & local conc.	0.751	0.107	5.44	reject	141.1
7	Translog	CRTS & local conc.					
a	setting *wr* wage ratio =1		0.753	0.106	5.43	reject	141.0
b	setting *wr* wage ratio = 0.6		0.741	0.11	5.59	reject	139.2

Note: in regressions 1 to 6 female labour units are converted to male labour units using the actual values of the female to male wage ratio (*wr*) observed in the sample villages. The average value of this ratio in this sample is 0.85 in Sample II.

6.7 Elasticities of substitution

A complete set of the estimates of the AES, the HEC and the factor demand elasticities for the TL specification with the linear labour composite function is given in Table 6.12. These are computed at the geometric mean of Sample II using the regression parameter estimates of Model C2 in Table 6.9. Ignoring the own-AES (values on the main diagonal) all other AES are positive, with one exception. The negative AES between bullock power and material inputs indicates these two inputs are *p*-complements. An increase in the price of bullock power reduces the input demand for material inputs. Amongst the set of positive AES, the AES between bullocks and effective labour is close to one; but all others are smaller than one. These estimates are quite different from previous estimates reported for Nepalese agriculture. For instance, in the set of AES reported for land, labour, bullock and fertilizers in Hamal (1991), all values are positive and higher than or close to one.[43]

The matrix of factor demand elasticities derived from the estimated AES is given in Panel (b) of Table 6.12.[44] The own-price ealsticities are all less than 0.5, with labour being the most price elastic. An own-price elasticity of effective labour of 0.463 seems to be on the high side but is within the range of estimates observed for other countries - Bapna *et al*. (1984) and Singh Squire and Strauss (1986c1).

The estimated HEC values reported in Panel (c) show that effective labour is a *q*-complement of land and material inputs. Increased usage of the latter two inputs increases the marginal product of labour. But human labour and animal labour are *q*-substitutes. Land is a *q*-complement for all other inputs, which is to be expected. In comparison to the HEC's computed for Sample I in Table 6.7, there is a reversal in sign in the HEC between effective human labour and material inputs. The positive HECS between human labour and material inputs in Table 6.12 is a more reasonable result than the negative HEC in Table 6.7. In both tables, however, the anomalous result of a negative HEC between bullock power and material inputs remains.

[43] Hamal's estimates are not specific to the *tarai* region agriculture, but for all of Nepalese agriculture, using a time series profit function estimation technique with aggregate data (Hamal 1991:142).

[44] The factor demand elasticities (ε_{ij}) are readily derived from the AES since $\varepsilon_{ij} = \sigma_{ij} \cdot s_j$, where s_j is the share of input *j* in the total value of production (Sato and Koizumi 1973).

Table 6.12 Elasticities of Substitution and Input Demand
(with $Le = F + \theta H$ effective labour composite)

(a) Allen partial Elasticities of Substitution

	Effec. Lab.	Effec. Land	Bullocks	Materials
Effective Lab.	-1.415	0.663	1.076	0.533
Effective Land		-0.493	0.173	0.092
Bullocks			-4.801	-0.662
Materials				-1.523

(b) Factor Demand Elasticities

	Effec. Lab.	Effec. Land	Bullocks	Materials
Effective Labour	-0.463	0.322	0.081	0.059
Efective Land	0.217	-0.239	0.013	0.01
Bullocks	0.352	0.084	-0.363	-0.074
Materials	0.174	0.045	-0.05	-0.169

(c) Hicksian Elasticity of Complementarity

	Effec. Lab.	Effec. Land	Bullocks	Materials
Effective Labour	-2.541	1.921	-2.117	0.502
Effective Land		-1.947	3.133	0.729
Bullocks			-6.549	-3.044
Materials				-2.605

Source: Computed from parameter estimates of Table 6.9 at the mean of the data.

6.8 Summary

The production function estimation results in this Chapter constitute the first step in the sequential estimation strategy of a farm household model that allows for labour heterogeneity. This first step estimation had two objectives: to test for heterogeneity between family and hired labour, and to generate a complete set of the factor demand elasticities and elasticities of substitution to describe the production technology.

The test for labour heterogeneity in itself had two components: a test of the separability of the labour inputs in the production function, and a test for the preferred specification of an aggregator function that converts family and hired labour into a composite labour variable, measured in effective units. The test for the separability of hired and family labour was based on a translog function specified with family and hired labour as distinct inputs, using a sample of households that used both family and hired labour inputs in crop production. The results showed that family and hired labour are only weakly separable from the other three main inputs-land, bullock power and material inputs. More restrictive types of labour separability are not accepted; and in particular the Cobb-Douglas form is strongly rejected.

The parametric estimates with family and hired labour as distinct inputs show these two labour categories are very close substitutes in both a price and quantity sense. At the sample mean the estimated partial Allen elasticity of substitution between family and hired labour is positive and large (greater than 10). The Hicksian elasticity of complementarity is negative. The negative HEC implies that the marginal product of one type of labour is reduced through increased application of the other.

The labour separability results make it feasible to embed the test for labour heterogeneity in a production function specification that compares alternative ways of aggregating family and hired labour into a composite effective labour (Le) input. Five different effective labour functions, allowing for different values of the elasticity of substitution and the marginal rates of substitution between family and hired labour, were estimated using a nested production function structure. The estimation was done over a larger sample of households that includes those cases not reporting any use of hired labour. The preferred effective labour aggregator function was the linear composite of the form $Le = F + \theta H$. The estimated value of θ was 0.751 with a standard error of 0.107. This estimate of θ is significantly less than one, implying that, although family and hired labour are perfect substitutes, they are not equivalent in efficiency units. There is a constant difference in their efficiency, with family labour being more productive.

The robustness of the result that θ is significantly less than one was checked through sensitivity analyses, allowing for alternative functional specification and parametric restrictions, as well as by varying the rate at which female work-days was converted into equivalent male work-days. These changes affect the estimate of θ only slightly and without negating the result that θ is significantly less than one.

The finding of a higher productivity of family labour in own farm crop production in the *tarai* region on Nepal is consistent with the results of Foster (1994) and Deolalikar and Vijverberg (1987) with Indian and Malaysian data. Both of these previous studies, however, found evidence for imperfect substitution between family and hired labour which is a result that differs from this study.

The complete set of the estimated production function parameters (Table 6.9) and the matrix of the various elasticities of input substitution derived from the estimated parameters (Table 6.11) are quite reasonable. The estimated AES are all less than one indicating again that the simpler Cobb-Douglas specification is inappropriate for this data set. The estimated HEC values show effective labour is a q-complement for land and material inputs but a q-substitute for bullock power.

Finally, this Chapter also shows there is considerable scope for functional mis-specification in treating different types of labour as distinct inputs in situations where the data contains many cases with a zero value for particular labour categories. (See Appendix 6.1). A nested production function structure with an effective labour variable appears to be superior to an estimation procedure that treats family and hired labour as distinct inputs.

Appendix 6.1

The Translog Specification with Family and Hired Labour as Distinct Inputs

This appendix presents the translog production function estimation results with family and hired labour as distinct inputs, using Sample II which contains a large proportion of households that do not use any hired labour. A conventional approach to resolving the problem of zero values of some inputs in the translog specification has to been to convert all zero values to one (or another small positive number).[45] An adjustment of this type seems particularly unsuitable with the data in Sample II because almost 60% of the households will have a zero level of hired labour input. This presumption is clearly borne out by the estimation results given in Appendix Table 6.13 for the Translog and Cobb-Douglas specification.[46]

The parametric estimates, especially for the labour inputs, in Appendix Table 6.13 vary greatly from the specification that uses the Le = F + θH linear composite for effective labour or even the homogenous aggregator. In the Cobb-Douglas estimates of Model 6 in Appendix Table 6.13, the output elasticity with respect to family labour (α_1) is less than 0.06. The elasticity parameter with respect to hired labour is less by a factor of ten, and is not significantly different from zero. The effect of these parameter estimates is that the marginal product of family labour at the mean of the data is now estimated to be only Rs. 3.05 per male work day while that for hired labour is Rs. 0.59 per day. These are very unreasonable estimates, given that wage rates are about Rs. 9.9 per male work-day and differ markedly from the marginal products computed with the homogeneous and linear composites in Table 6.10.

The translog specification in Model 5 gives more reasonable results than the Cobb-Douglas but anomalies remain. The elasticity with respect to hired labour is (0.032) is considerably smaller than compared to family labour (0.258) at the mean of the data. At this point the estimated marginal product of family labour (Rs. 13.86) is

---

[45] Jacoby (1993) is one example of this commonly used *ad hoc* procedure.

[46] The prior restriction of constant returns to scale is imposed in both models. In the translog specification (Model 5) the data are scaled to the geometric mean which results from the conversion of the zero values of hired labour to one; and weak separability and local concavity are also imposed.

reasonable, but the marginal product of hired labour (Rs. 3.11) is considerably underestimated. In the translog the production elasticity parameters (α_1 and α_2) are sensitive to the point of scaling of the sample data. If the hired labour variable is scaled to the mean, not of Sample II with the arbitrary conversion of zero values to one, but to the mean of Sample I (which includes only those households with positive inputs of both family and hired labour), then the estimated marginal products become reasonable. At the mean of the Sample I data, the estimated marginal product of family labour is Rs.14.46 and of hired labour Rs.10.97 using the translog specification. While the production technology of the farms which employ both types of labour appear to be well represented by this specification it does not reflect accurately the common technology of the pre-dominant group of farmers who do not use any hired labour.

Note that despite the unreasonable parameter estimates with family and hired labour as distinct inputs, the goodness of fit criteria for Models 5 and 6 are very close to that for the specification with a composite effective labour function. In Model C2 in Table 6.9 the residual sum of squares is 23.71 and in the translog specification of Appendix Table 6.23, the residual sum of squares increases only marginally to 24.1. There is no change in the R-Sq. between observed and predicted values. It appears the substantial model mis-specification with family and hired labour treated as distinct inputs, with zero values converted to one's, could be masked by the standard goodness of fit criteria, especially in a non-linear estimation procedure.[47]

The estimation results in Appendix Table 6.13 clearly illustrate the problem that can occur when the *ad hoc* procedure for converting zero values in situations when such zero values occur in a large proportion of the sample. The results with the effective labour aggregator function clearly are more economically meaningful even though statistically the model fit may not be very different.

[47] It may be that when the estimation can be done by OLS the wider variety of mis-specification diagnostics tests (such as the RESET) could adequately differentiate Models 5 and 6 in Appendix Table 6.13 and models with effective aggregator functions. These standard specification tests are not applicable for non-linear estimation (Greene 1993), and so have not been applied here to discriminate in a more formal way between the specification of Model 5 in Appendix Table 6.13 and Model C2.

Appendix Table 6.13

Translog Production Function with Family and Hired Labour as Independent Inputs

Dependent Variable	: Log real composite farm output
Sample subset	: Sample II (households using some family labour) (N =679)
Prior Restrictions	: Constant returns to scale
Estimation Method	: Non-Linear least squares with heteroskedasticity corrected errors

		Model 5		Model 6	
additional restrictions:		WS* and local concavity		Cobb-Douglas	
Variable		*Coeffic.*	*t- ratio*	*Coeffic.*	*t- ratio*
Family lab. (F)	α1	0.258	*8.14*	0.0567	*4.33*
Hired Lab (H)	α2	0.0328	*4.75*	0.0054	*0.87*
Efective Land (Ae)	α3	0.501	*10.37*	0.753	*24.90*
Bullocks (B)	α4	0.093	*4.44*	0.055	*3.32*
Material inputs (M)	α5	0.115	*4.47*	0.129	*5.52*
Educ of Head		0.006	*1.12*	0.0045	*0.74*
Big Farm dummy		0.003	*0.11*	-0.014	*0.63*
Second order coefficients (x 10)					
FxF	α11	*0.475*	*5.70*		
HxH	α22	*0.245*	*4.61*		
FxH	α12	*-2.110*	*3.58*		
AexAe	α33	*-1.290*	*1.57*		
BxB	α44	*0.360*	*2.41*		
MxM	α55	*0.715*	*1.70*		
FxAe	α13	*0.590*	*2.00*		
HxAe	α23	*0.070*	*1.89*		
FxB	α14	*-0.510*	*2.52*		
HxB	α24	*-0.060*	*2.36*		
FxM	α15	*-0.357*	*1.69*		
HxM	α25	*-0.045*	*1.63*		
AexB	α34	*0.569*	*1.68*		
AexM	α35	*0.047*	*0.98*		
BxM	α45	*-0.359*	*1.33*		
R -Sq. between observed and predicted		0.967		0.963	
Residual sum of Squares		24.1		26.3	
Log Likelihood		133.2		105.6	
Test for Cobb-Douglas form :					
Wald test statistic (with 7 DF)		60.00	(reject)		
Estimated labour marginal product at mean (Rs.)					
family labour		13.86		3.05	
hired labour		3.61		0.59	

*WS = weak separability of the labour inputs

CHAPTER VII
LABOUR SUPPLY ESTIMATION

7.1 Introduction and Motivation

The results in Chapter VI of a production-function-based test for the heterogeneity of family and hired labour as inputs in Nepalese *tarai* region agriculture indicated that family and hired labour are perfect substitutes in farm production but with different productivity. The production function estimation was based on choosing among alternative functional forms for creating an effective labour (Le) composite function, which aggregated family (F) and hired labour (H) inputs. The preferred aggregator function was a linear composite, $Le = F + \theta H$. The estimated value of θ was 0.75 which was shown to be significantly less than one. When measured in *effective* units, one unit of hired labour is equal to 0.75 units of family labour.[1]

The production function based result on labour heterogeneity has important implications for the appropriate methodology as well as model specification in estimating the labour supply component of the farm household model. As discussed in Chapter III, the main methodological implication is that the linear form of the labour heterogeneity still makes the farm household model *recursive* in its production and consumption components. The constant efficiency difference between hired labour and family labour implies a farm household model structure wherein the effective wage rates faced by household members will differ according to the labour market exposure of the household in the hired labour market - i.e. whether the household is a net buyer or net seller of labour. The efficiency difference, however, is independent of the levels of labour and other inputs used on the farm since θ is a constant. Hence, the "effective" wage rate faced by family labour applied to the farm is still parametrically given to the household. Consequently, the labour supply estimation can be done separately from the

[1] As noted in Chapter I, this thesis does not delve into alternative explanations for the higher productivity of family labour. One common justification is that the effort applied per unit of time is likely to be lower for hired than for family labour (Feder 1985). Another reason could be that family labour acquires some farm-specific experience (Rosenzweig and Wolpin 1985).

production side of the model, but with the necessary adjustment to the observed market wage rates for the difference in productivity represented by θ.

This Chapter presents the estimation procedure and regression results for the labour supply behavior of farm household members taking into account the observed heterogeneity between family and hired labour in farm production. The resulting set of estimates of the parameters of the labour supply function completes the description of the behavioural response of the farm household. However, another equally important motivation for this Chapter is to provide independent corroboration of the result derived in Chapter VI that family and hired labour are heterogeneous inputs in farm production.

The production function based test for $\theta < 1$ is robust with respect to several alternative functional specifications and parametric restrictions (Table 6.5). While this is strong evidence for labour heterogeneity, it is not conclusive in the sense that it does not rule out alternative explanations of why the estimated θ could be less than one in the production function regression results. The main concerns are of data aggregation bias and unobserved quality differences in the family and hired labour inputs that are independent of their family or hired status.

The farm management data utilized in the regression analyses of Chapter VI is given only at the aggregate household level in the MPBHS. Individual characteristics of persons who supply the family and hired labour on a particular farm are not observed. There could be some unobservable quality differences between family and hired labour that give rise to a difference in their marginal productivity independently of the distinction between family and hired labour categories. For instance, in a particular household family labour may consist solely of prime age workers, while the hired labour input may be of young teenagers or older adults. Similarly education levels of family and hired labour may be different. Another source of unknown bias in the estimate of θ may arise from the aggregated nature of the farm level production function estimated in Chapter VI, since labour requirement and productivity will differ across crops. Therefore it is important to find additional evidence for the lower marginal productivity of hired labour independently of the production function estimation. The labour supply regressions presented in this Chapter provide a mechanism for such an independent verification.

It is feasible to test whether the observed labour supply behaviour is consistent with the heterogeneity between family and hired labour detected in the production function estimation by comparing alternative labour supply model specifications. A labour supply model that equates the opportunity cost of family labour in all households to the observed market wage rate for off-farm work is consistent with family and hired labour being homogeneous production inputs. Model specifications using effective wage rates (based on θ) that vary according to the labour market exposure of the household are consistent with labour heterogeneity. Standard diagnostics for model selection can be used to check whether the common wage labour supply model performs better than the varying effective wage model. If these model selection tests find in favour of the varying wage model, this result can be interpreted as conforming with the productivity differences in family and hired labour inputs in farm cultivation. [2]

The superior performance of a labour supply model with varying effective wage rates does not necessarily prove there has to be a difference in efficiency between family and hired labour in the production function. Such a result could also be consistent with other explanations which lead to differences in the effective wage rate for family labour to work on its own farm and the wage rate applicable for work on the hired labour market. A "wage gap" of this type could arise from differences in tastes between working on one's own farm and working on the hired labour market (Lopez 1984), fixed costs to seeking outside work (Cogan 1981), etc. The tests for the labour supply model specifications are not designed to discriminate between alternative sources or explanations for varying effective wages. They only

[2] The only direct connection between the labour supply estimation and the production function estimation is that the actual value of the θ parameter estimated in the latter is used to create the effective wage and non labour income variables for the labour supply regressions. Since the labour supply data module has not been utilized in generating the estimate of θ, this is an independent data set that can be checked for *conforming* evidence of θ being less than one, if not necessarily *confirming* that there is a genuine difference in the efficiency of family and hired labour.

discriminate between models relying on a common market wage and models that specify varying effective wage rates, based on whether households are net buyers or sellers of labour. A result that labour supply behaviour conforms with the evidence in the production function for labour heterogeneity increases the likelihood that there is a genuine efficiency difference between family and hired labour. If the results reported in Chapter VI were due solely to other unobserved factors not taken into account in the production function estimations, it would be an unlikely co-incidence if these extraneous factors also led to a labour supply model characterized by higher effective wage rates for family members on farms that employ hired labour.[3]

A secondary objective of this chapter is to obtain accurate estimates of the parameters of interest of the labour supply functions (i.e., wage and income elasticities for various household types) by correctly specifying the appropriate effective wage that correctly defines the leisure-labour equilibrium faced by different households. It is of some interest to check how these elasticities will differ between specifications that recognize labour heterogeneity (and hence the resulting wage-gap) and those which treat both types of labour as homogeneous inputs.

In the following sections, the next (7.2) briefly discusses how the household's optimal labour supply/leisure demand conditions are affected by labour heterogeneity; and, in particular, how the effective wage rates at the equilibrium are related to the θ parameter. Section 7.3 provides a brief summary of the individual and household level data used in the labour supply function estimations. Section 7.4 discusses the specifications of the alternative models and the related model identification issues. Section 7.5 provides the labour supply regression results for male household members, including model selection tests and estimates of the wage and income elasticities of labour supply from the alternative models. Section 7.6 gives a similar set of results for the labour supply of female household members. Section 7.7 concludes the chapter.

[3] The precise form of the wag gap created by the θ efficiency difference factor is illustrated in Figure 7.1 in Section 7.2. The labour supply model specification tests reported in this Chapter cannot discriminate between alternative underlying explanations that give rise to Figure 7.1. But they can rule out other explanations that result in an arbitrary wage gap model inconsistent with Figure 7.1.

7.2 Labour Supply Implications of Linear Heterogeneity

The effect of the linear aggregator function for effective labour ($Le = F + \theta H$ with $\theta < 1$) is that one unit of family labour when applied to the family farm can substitute for $1/\theta$ units of hired labour, without affecting output. Let w denote the market wage rate paid per unit of hired labour. For the farm household which is hiring in labour, the effective wage rate that can be applied to family labour devoted to the family farm is w/θ (because one unit of F substitutes for $1/\theta$ units of H for which the hired wage cost is w/θ). Moreover, since the $1/\theta$ conversion factor is independent of the actual levels of family and hired labour (and other inputs) applied on the farm, there is a constant difference in the marginal product of hired labour *vis a vis* family labour. The marginal rate of substitution between family and hired labour (which is the ratio of the two marginal products) is constant for all households. The first order conditions for the optimal levels of family and hired labour inputs can be derived as if the household faced a wage rate of w for hired labour and an internal wage rate of w/θ for family labour when applied to its own farm (see Section 3.5 in Chapter III). Even though w/θ is a shadow wage rate which applies only to an artificial internal household labour market for family labour, the family labour supply behaviour can be modeled as if the household were to take the w/θ wage as parametrically given to the household (when the market wage rate w can be assumed to be exogenous).

This maintains the recursive nature of the farm household model whereby labour supply decisions can be modeled independently of other production input choices. Consequently, the traditional estimation strategy can be followed, where the production function and labour supply/consumer demand systems are estimated separately, with a minor adjustment to market wage rates to derive the effective wage rates that represent the consumer-household's labour/leisure equilibrium.

A second major implication is that the efficiency difference between family and hired labour as production inputs affects the effective wage rate, and hence labour supply decisions, only of households who hire in labour (or are at the margin of deciding to employ the first unit of hired labour). For a small farm household that supplies family labour on its own farm and works as well on the off-farm hired labour market at a wage rate of w, the optimal family labour input on own farm cultivation is conditioned by the market wage rate, w and is not affected by the θ parameter. Hence, the effect of the efficiency difference between family and hired labour as production inputs on the labour supply behaviour of farm households depends on the net labour market position of the household.

There are three mutually exclusive household categories based on the net labour market position. If M represents the amount of family labour supplied on the hired labour market (at wage w) and H represents the amount of labour hired in (again at wage w), the three mutually exclusive household categories are:

Category 1: family labour hired out ($M > 0$) and no labour hired in ($H = 0$). This category represents landless and small farmers who are net sellers of labour.

Category 2: family labour not hired out ($M = 0$) and extra labour hired in ($H > 0$). This category represents big farmers who work on their own farms and also hire in labour.

Category 3 family labour not hired out ($M = 0$) nor extra labour hired in ($H = 0$). This category represents "autarchic" households who equate labour demand and supply on the family farm solely from family sources.

The possibility of both $M > 0$ and $H > 0$ is explicitly ruled out if $\theta < 1$ and the wage rate for hiring in and hiring out are the same (as assumed in this study). Since one unit of family labour is equal to $1/\theta$ units of hired labour, if the two wage rates are the same, the household always gains by transferring its labour from market wage work (activity M) to own farm work (activity F) so that it can reduce hired labour demand by $1/\theta$ units for every unit of labour so transferred.[4]

[4] The theoretical inconsistency of both M and H being positive for a household ignores seasonal variation in labour supply and demand and changes in the net labour market exposure of a household.

The prediction that, in the presence of labour heterogeneity with family labour being more productive than hired labour, farm-households should not simultaneously hire in and hire out labour is a strong result. It can be readily tested with the sample data of this study. Table 7.1 bears out that the *tarai* region sample households indeed do not report simultaneous hiring in and hiring out of labour. Out of 686 own-farm operator households for which a matched set of labour demand and labour supply data could be assembled for the subset of the MPBHS sample used in this study, only 22 households - about 3% - report using hired labour as well as some member of that household working on the off-farm labour market. Most cases of both hiring in and hiring out are found in the smallest land size category, where the amounts of hired labour used are very small. Out of 199 big and medium farm size households only three report simultaneous hiring in and hiring out labour.

Table 7.1 Market Labour Supply Exposure of Sample Households

Operated Land Sampling Strata	Sample Size (N)	Total N with matched data	Number of Households reporting : Hiring IN labour	Hiring OUT labor	Hiring IN and OUT
col. 1	2	3	4	5	6
BIG	103	77	69	2	1
MEDIUM	123	122	87	10	2
SMALL	227	214	89	62	8
MARGINAL	281	273	48	179	11
ALL cultivators	734	686	293	253	22 (3% of col. 3 total)
LANDLESS	273	240	0	240	0
TOTAL	1007	926	293	493	22

Note : The total N reported in column 3 is the number of households for which both the labour supply and labour demand data could be matched from the different parts of the MPHBS data records. Labour supply records for some landless households are missing.

Given that the survey data refers to the entire annual cropping cycle, and given the very time specific nature of agricultural operations and strict gender-related division of labour, this is a striking result, which is consistent with the type of labour heterogeneity detected in the production function estimation in Chapter VI.

The equilibrium conditions for labour demand and supply for each of these three household categories were illustrated in Figure 3.4 in Chapter III for the case where θ^* (the marginal rate of transformation between family and hired labour at the optimum labour allocation) was less than one. The theoretical derivation in Chapter III was based on a general form of labour heterogeneity in a production function with a separable labour nest. The specific linear composite indicated by the empirical results of Chapter VI further simplifies the theoretical structure and estimation strategy. It is not necessary to compute θ^* at the optimum labour allocation for each sample household to define the effective wage rate applicable for the labour supply equilibrium. The applicable effective wage rate can be simply determined from the constant θ parameter and the observed market wage rate.[5]

Figure 3.4 makes it clear the wage rates which define the equilibrium labour supply position on the labour supply curve are different for the three household types. For Category 1 (landless and small farm) households, where total labour supply is greater than own-farm demand (M > 0), both the implicit valuation of family labour and the marginal return to own labour in farm production are equal to the wage rate (w) received on the hired labour market (Fig 3.4a). This result is independent of the value of θ. At the optimum labour allocation, the real cost of labour in all household activities is equalized to the market wage rate.

For Category 2 households (big farms where labour is hired in) the optimal use of hired labour equates the marginal product of hired labour to the hiring-in wage rate (w); whereas the allocation of family labour to own farm production equates the marginal returns to w/θ (Fig. 3.4c).

For Category 3 (autarchic) households which neither hire in nor hire out any labour the equilibrium conditions set the effective wage rate for the labour supply equal to

[5] As indicated in Eq. 3. 20, in the general case θ^* is a function of the optimum levels of F and H. In the case of the linear composite for effective labour θ^* becomes a constant, equal to the estimated value of θ (0.75) in Chapter VI for all households, and this value is independent of all input levels.

the marginal product of labour on the family farm. This common value for the real cost of labour and the marginal product (ω) lies between w and w/θ (Fig.3.4b); and this shadow wage rate (ω) can vary from household to household, even if they face the same market wage rate. For autarchic households, a small change in the market wage rate will have no effect on equilibrium labour supply or demand as long as the autarchic nature of the household is preserved. That is, if the wage change is not large enough to make the autarchic household change its status into a Category 1 or Category 2 household. The underlying farm-household model is not separable for autarchic households even with a linear form of labour heterogeneity.

The effect of the lower efficiency of hired labour on the own-farm labour allocation decision of the three different categories of farm households can be summarized in terms of the wage at which an unit if effective labour is available to each household type. This is illustrated in Figure 7.1, where labour is measured in effective units on the horizontal axis. As in figure 3.4, the VV curve denotes the marginal rate of substitution between leisure and consumption for a representative individual worker in a farm household. The stepped bold line wSS'S" represents the supply price at which an effective unit of labour is available to the three different farm households. The classification of the labour market exposure of the farm household is based on the amount of labour required for own-farm cultivation which determines where the demand for effective labour (or the marginal product of effective labour) schedule intersects with the supply curve.

As drawn, Y^S reflects the demand for effective labour on a small farm that has extra labour to sell on the off-farm labour market at wage w. For such a household, the price at which a unit of effective labour is available for own farm cultivation is also w (since units of family and effective labour are equivalent). When the farm size is large enough to require hired labour for cultivation, the supply price for a unit of effective labour becomes w/θ. Such a farm is represented by the Y^B schedule for the demand for effective labour. The large farm faces an effective wage rate of w/θ because a standard unit of hired labour at wage w represents only θ units of effective labour. The intersection of the marginal product of labour curve denoted as Y^A with the upward sloping section SS' of the supply curve represents an autarchic household that neither sells nor buys any labour. The marginal product of family labour, and hence its shadow price, for the autarchic household is between w and w/θ. Figure 7.1 indicates that if preferences for leisure and production technology are constant for all households, the sole basis for the classification of households according to their exposure in the hired labour market will be farm size, which determines the horizontal distance at which the marginal product curve intersects the supply curve.

Figure 7.1 Supply Price of Agricultural Labour with Linear Heterogeneity

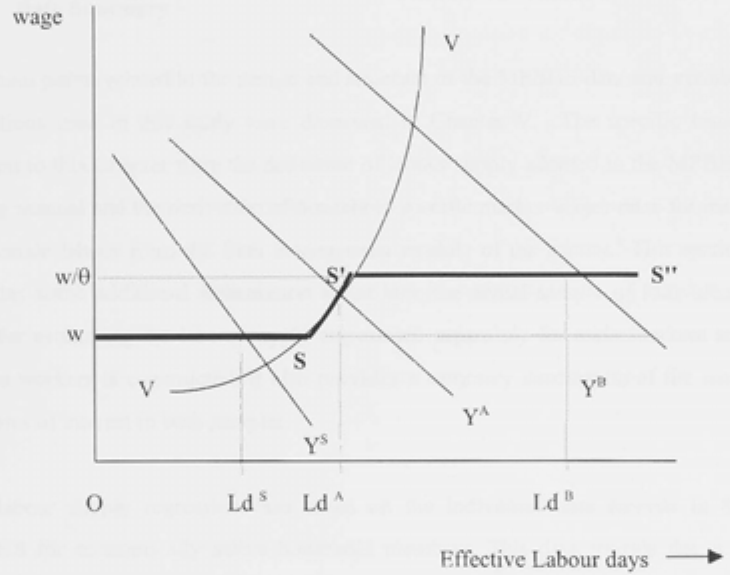


Figure 7.1 also clearly shows that the effect of the θ factor in the difference in productivity of family and hired labour is equivalent to a wage gap in the price at which an effective unit of labour is available to different farm size categories.[6]

[6] There is a straightforward equivalence between the wage gap in Figure 7.1 and the price gap created by the difference between c.i.f. import and f.o.b. export prices in a model of international trade with transport costs. The small farmer who sells labour in the off- farm market is exactly analogous to an exporting country where the domestic supply price is the f.o.b. price. The big farmer who "imports" hired labour is equivalent to an importing country that has a domestic supply price given by the higher c.i.f. price. In between is an autarchic country that does not trade if its domestic supply price is in between the f.o.b. and the c.i.f. price. (See Dixit and Norman 1980).

The per hectare labour input in effective units applied to these three different farm categories will vary because the wage rate that each of the three household category faces is different. There is a rising supply price of labour only among the autarchic households. For the small and big farm households, although there is a gap in the effective wage they face, the supply of effective labour is available at a fixed price. Within the set of big farm households that rely on hired labour, there is no further difference in the wage cost: all big farms face the common wage of w/θ for one unit of effective labour.[7]

7.3 Data Summary

The main points related to the design and structure of the MPBHS data and variable definitions used in this study were discussed in Chapter V. The specific issues relevant to this Chapter were the definition of labour supply adopted in the MPBHS coding manual and the derivation of household specific market wages rates for male and female labour from the farm management module of the survey.[8] This section provides some additional information about how the actual sample of individuals used for estimating the labour supply regressions separately for male workers and female workers is constructed. It also provides a summary description of the main variables of interest in both samples.

The labour supply regressions are based on the individual case records in the MPBHS for economically active household members. This data records the total days of work in different categories for each of two six-monthly cropping season survey cycles. The full MPHBS data contains employment details for all economically active persons aged 10 or above. The sample used for the labour supply regressions in this Chapter is further restricted in three ways to individuals who are: (i) aged 15 to 60, (ii) are in a familial relationship with the household head, and (iii) whose usual occupation codes are related to farm cultivation.

[7] This assumes there are no quantity constraints to the supply of hired labour at the market wage, w.

[8] See Section 5.3(b) in Chapter V for details.

The age group restriction is imposed to ignore the occasional labour supply records for children aged 10 to 15 and the elderly. Also excluded are individuals not related to the household head. One important category thus excluded is the group of domestic and farm servants reported to reside in the household of their employer. The labour supply for such individuals would differ from that of family members in the household. The contractual arrangements for permanent or semi-permanent farm labourers can be complex, and their labour supply behaviour need not be related to an observed market wage rate for hired casual labourers.[9] The final occupational code restrictions are imposed because the effective wage rates for family labour are to be created from the observed wage rate for hired labour in farm cultivation. This wage rate need not be a good proxy for the marginal returns to extra work in other village occupations - shopkeepers and caste-based professional occupations such as shoemakers, blacksmiths, etc.[10]

The final adjusted sample size available for the labour supply regressions consists of 2,542 person-season records for male family members and 2,288 person-season records for female family members. Given that both summer and winter season records are available for most individuals, this represents a sample size of about 2,500 economically active family members, from the 1,007 sample households, in the specific labour force categories chosen above.

In the case of female family members, data was also compiled for those individuals aged 15 to 60 who are reported as economically inactive in order to generate a sample which includes active and inactive family members to counter potential sample selection bias in the regression results. A total of 481 women aged 15 to 60

[9] Bardhan and Rudra (1981) provide details of the terms of various forms of "attached" labour contracts in West Bengal. A similar variety of contractual modes also occurs in the Nepal *tarai* region for which data was not collected in the MPBHS. Hence the labour supply of farm servants cannot be separately modeled in this study.

[10] The specific occupation codes used to restrict the sample to agricultural workers from the detailed occupation coding list of the MPHBS were codes 61 (farmers) and 62 (agricultural and animal husbandry workers) and 99 (common labourers). The last category would include some non-agricultural labourers such as porters and *coolies*, but their wage rates are likely to be closely related to the daily wage rate for hired agricultural labourer.

were reported to be economically inactive. When compared to the final sample of about 1200 economically active women aged 15 to 60, this represents a non-participation rate of about 28%. There is no corresponding problem of potential sample selection bias in the sample of male family members. Only 26 cases of economically inactive men aged 15-60 were recorded in the entire sample.[11]

A summary description of the main variables used in the labour supply regressions for both the male and female samples is given in Table 7.2. It also provides a breakdown of the data for the main variables by the three categories of households based on their labour market exposure. Table 7.2 clearly shows the average number of days worked is strongly related to labour market exposure. In both male and female samples, average work-days is highest for Category 1 households (labour hire out) followed by Category 2 (labour hire in) and lowest in the autarchic (Category 3) households. The association between labour market exposure status and average farm size is also evident. Individuals in the labour hiring out sample come from landless and small farmer households, with an average farm size of 0.88 hectares. Individuals in the autarchic category have an intermediate average farm size of 2.2 hectares, while average farm size in the labour hiring in sample is 4.67 hectares. There is a similar relationship between per capita non-labour income and the labour market exposure categories, since this income category includes the imputed returns from land ownership (farm profit). Table 7.2 reveals a large variation in the non-labour income variable. This helps to accurately estimate the income effect on labour supply. The sample variation on real wage rates is also high, reflecting regional patterns. For instance, in the sample of male workers the real wage rate ranges from 2.63 to 7.3 (kilograms of paddy a day).[12]

[11] This is a very small proportion of the final sample of almost 1300 active male individuals used for the male labour supply regressions.

[12] The entries in Table 7.2 differ slightly for the male and female samples, since these means are un-weighted averages summed over working individuals not households.

Table 7.2: Labour Supply Data Summary for Male and Female Family Workers

Variables	Mean	Std Dev.	Min.	Max.
Male Workers Sample (N = 2542)				
Total labour supply # (all categories)	88.6	41.7	0	177
Category 1 Households (hire out) (n = 1240)	108.5	38.7	0	177
Category 2 Households (hire in) (n = 813)	78.1	36.3	2.2	173
Category 3 Households (autarchic) (n = 489)	60.4	33.9	5	171
Real Household non-labour income ##	5511	6318	27	46,180
Per capita real non-labour income ##	619	629.6	7	5,905
Category 1 Households (hire out labour)	255	242.1	7	2,359
Category 2 Households (hire in labour)	1123	802.2	152	5,905
Category 3 Households (autarchic)	681	294.6	145	2,543
Average market wage rate for hired labour				
Real female wage rate (Kg. of paddy/day)	4.19	0.74	2.63	7.43
Real male wage rate (")	4.51	1.06	2.65	8.19
Age	34.9	12.9	15	60
Family size	8.56	5.2	1	28
Number of adult male workers per household	2.32	1.37	1	7
Number of adult female workers per household				
Total farm cultivated area	2.52	3.01	0	33
Category 1 Households (hire out labour)	0.88	1.39	0	9
Category 2 Households (hire in labour)	4.67	3.79	1.47	33
Category 3 Households (autarchic)	2.98	2.47	1.2	13
Female Workers Sample (N = 2288)				
Total labour supply # (all categories)	56.5	35.4	0	180
Category 1 Households (hire out) (n = 1103)	68.0	38.8	0	174
Category 2 Households (hire in) (n = 724)	51.4	28.8	0	180
Category 3 Households (autarchic) (n = 461)	45.5	31.1	2	170
Real Household non-labour income ##	5471	5932	27	46,180
Per capita real non-labour income ##	575	494.2	7	3,769
Category 1 Households (hire out labour)	252	223.1	7	1,448
Category 2 Households (hire in labour)	994	543.7	76	3,769
Category 3 Households (autarchic)	729	389.1	108	2,542
Average market wage rate for hired labour				
Real female wage rate (Kg. of paddy/day)	4.22	0.72	2.63	7
Real male wage rate (")	4.57	1.07	2.65	8
Age	31.9	11.7	15	60
Family size	8.64	5.4	1	28
Number of adult male workers per household				
Number of adult female workers per household				
Number of children aged 0-5	2.06	1.77	0	9
Number of children aged 6-9	0.8	0.9	0	6
Economically Inactive female individuals: (n = 481)				
Per capita real non-labour income	704.1	931	90	5,905
Family size	8.56	5.2	1	28
Number of adult male workers per household	2.03	1.25	1	7
Number of adult female workers per household	2.69	1.39	1	10
Age	34.1	14.9	15	60

This is the average number of work-days in each of two 6 monthly cropping season.
Nominal income (in Rupees) deflated by the village price of paddy per kg.

7.4 Model Specification and Identification

From Eq. 4.11 (in Section 4.3 of Chapter IV) the specification of the labour supply equation in the second step of the farm household model estimation strategy, based on the prior estimate of θ obtained in the first step, is given by

$$(7.1) \quad Ls = \ell\,(\,w^*(\theta),\ \pi(\theta) + E,\ \mathbf{S}, \beta) + e$$

where $w^*(\theta) = w$ for Category 1 households

$= w/\theta$ for Category 2 households

$= \omega$ where $w \leq \omega \leq w/\theta$ for Category 3 households

$\pi(\theta) + E$ = non-labour income, including imputed farm profits.

β = the labour supply parameter set.

For the empirical implementation of Eq. 7.1 a linear labour supply function is specified where the total labour days worked by an individual is regressed on the real effective wage rate, real per-capita non-labour income, and a set of individual and household level characteristics. The linear labour supply is chosen partly for the sake of simplicity and partly to aid in model identification. Because of the multiplicative form in which the θ factor affects the effective wage rate, a log-linear form cannot be used since the wage adjustment is converted into an intercept effect.

The labour supply estimation is done separately for male and female household members, allowing for cross wage effects. The general estimating equation is:

$$(7.2) \quad Ls_{ikh} = \alpha + \beta_1 w^{*m}_{kh} + \beta_2 w^{*f}_{kh} + \gamma\, PNLY_h + \tau'\mathbf{S} + e_{ikh}$$

Ls_{ikh} = total work days reported for an economically active member i in household h for crop season k (k = summer or winter).[13]

$PNLY_h$ = real per capita household non-labour income for household h

[13] The sample of individuals included in the labour supply regressions is further restricted to family members aged 15 to 60 who report their main occupation as being agricultural workers or own account farm operators.

S = a vector of individual and household characteristics which includes family size and composition, age and education level of the individual family relationship dummies, etc.

w^{*m}_{kh}, w^{*f}_{kh} = effective real male and female wage rates for household h in season k, which depend on the labour market exposure of the household as follows (suppressing the kh subscripts)

$w^{*j} = w^j$ if household is a net seller of labour (Category 1)

$w^{*j} = w^j/\theta$ if household is a net buyer of labour (Category 2)

$w^{*j} = w^j/\lambda_j^q$ for autarchic households (Category 3)

(with $\theta \leq \lambda_j^q \leq 1$)

where j = m (for male) or f (for female)
and q indexes the number of autarchic sample households.

The effective wage rates also enter into the computation of non-labour income since the imputed cost of family labour applied to farm production must be deducted from gross farm profits.[14]

The market wage rate for hired agricultural labour computed from the MPHBS can be distinguished by three different categories: (i) by sample household (ii) by gender, and (iii) by each of two six-monthly cropping seasons (the summer and winter cycle) for which farm management data was collected separately. This structure for the wage data allows some variation even in market wage rates for households within a single sample village cluster. All households in a specific sample cluster are not imputed an average village specific wage. Where a particular household reports hiring labour for crop production in a specific season, the market wage rate for that household is computed as the average of the actual wages paid out during that cropping season. A village level average wage is then computed as the unweighted average of the hiring-in wage rate reported by all such households in the sample village. This average wage paid to hired labour is then imputed as the village level wage received by individuals in the small farm and landless households for

[14] See Section 5.3 in Chapter V. The variable PNLY in Equation 7.2 is just NLY in Equation 5.3 divided by family size and the village specific price of paddy. This converts a nominal household level variable to a real per-capita level variable.

whom no household level hired wage is reported directly in the MPHBS data.[15] All household members in a specific gender group get assigned the same wage rate since individual level wage rates are not reported in the survey.

Once the nominal household-level wage rates (for male and female members) are computed in this manner, they are deflated by a village-specific price of paddy. For simplicity, the real wage conversion is based only on the price of paddy rather than the price of a composite consumption good for which complex indexing procedures would have been necessary.[16]

The real market wage rates (w/p) derived in this manner are converted into real effective wage rates (w^*/p) based on the categorization of households by their net labour supply position. The description of a household's net labour supply position (Category 1 to 3) is based on data for the whole annual survey period and does not vary by cropping cycle. The Category 1 to 3 description also carries over to all individuals in the same household.[17] Apart from the gender distinction, all individuals in a household are assumed to face the same market and effective wage. Since individuals in each gender group are assumed to be perfect substitutes for each other in own farm production, and individuals across gender groups are also perfect substitutes with a productivity difference related to the ratio of the market wage rates, it should not matter which particular family member works on the farm and which one works on the hired labour market.

[15] Since rural labour markets are very localized, the wage reported to be paid out by a labour hiring household is likely to be the wage received by the neighboring households that work on the hired labour market. In such a setting the fixed costs and other transactions costs to working on the neighbour's farm will tend to be negligible. See Section 5.2 for a fuller discussion of this issue.

[16] In the *tarai* region of Nepal rice is the major food consumption item. In many areas of the Nepal *tarai* wages to agricultural labour are often paid out in kind in units of paddy. The wage reported in the MPHBS are in part derived from the monetary conversion of in-kind wages paid in paddy.

[17] For instance, consider a sample household with two economically active males and one female member, where only one of the male individuals reports working as a hired labourer. The market wage rate for male labour is assigned as the effective wage rate for the other male worker also. The market wage rate for female labour is assigned as the effective wage rate for the female member, even though the latter two have not themselves reported individual work in the hired labour market.

Given a prior estimate of θ and the reported market wage rates, the effective wages are well defined for households in Categories 1 and 3. But the effective wage rate for autarchic households in Category 2 is not defined because λ_i^q cannot be identified since it varies across households. It is not appropriate either to use the market wage rate, *w*, adjusted by some constant factor because small changes in *w* will have no effect on the labour supply of individuals in autarchic households. If the market wage rate is attributed to the labour supply equilibrium of autarchic households, the expected values of wage coefficients (β) are zero since $\frac{\partial Ls}{\partial w^i} = 0$ for individuals in Category 3 households.

Section 7.4 below presents some preliminary regression results based on a sample that includes the autarchic households to detect whether this condition is met. But the observations on the autarchic households are dropped in subsequent regressions.

<u>Alternative model specifications</u>

Focussing on individuals in Categories 1 and 3 only, four separate models can be defined. These differ on whether the common observed market wage or the varying effective wage rate is used; and whether or not slope and intercept dummies are allowed for the two household categories. These alternative models can be represented in the framework of Eq. 7.2 as follows (ignoring for the moment the subscripts and the distinction between male and female wage rates):

Model A: common market wage; common parameters

(7.3A) $Ls = \alpha + \beta w + \gamma PNLY + \ldots$ for Category 1 & 2

Model B: common market wage; varying parameters

(7.3B.1) $Ls = \alpha^1 + \beta^1 w + \gamma^1 PNLY(w) + \ldots$ for Category 1

(7.3B.2) $Ls = \alpha^2 + \beta^2 w + \gamma^2 PNLY(w) + \ldots$ for Category 2

Model C: different effective wage rates; common parameters

(7.3C.1) $Ls = \alpha + \beta w + \gamma PNLY(w/\theta) + \ldots$ for Category 1

(7.3C.2) $Ls = \alpha + \beta (w/\theta) + \gamma PNLY(w/\theta) + \ldots$ for Category 2

Model D: different effective wage rates; varying parameters

(7.3D.1) $Ls = \alpha^1 + \beta^1 w + \gamma^1 PNLY(w/\theta) + ...$ for Category 1

(7.3D.2) $Ls = \alpha^2 + \beta^2 (w/\theta) + \gamma^2 PNLY(w/\theta) + ...$ for Category 2

Note : PNLY(w) is PNLY derived from valuing the cost of family labour at wage w; and PNLY(w/θ) is derived using the effective wage rate w/θ.

All four model specifications above can be represented within the general framework of Equation 7.2 through the use of appropriate dummy variables for the intercept and slope terms. For Models B, C and D several different variants occur within each set, depending on the combinations of dummy variables used. For instance in Model C, while the wage slope parameter β is assumed to be constant across individuals in Category 1 and 2 households, the γ or the α parameter for the intercept may be allowed to differ for individuals in Category 1 and 2 households.

Within the general structure specified in Equations 7.3, Models A and B are nested; and so are C and D. The selection of the appropriate model within each choice set can be based on testing the coefficients on the appropriate dummy variables that give rise to differences in the parameter estimates for Categories 1 and 2. The main interest in this Chapter, however, is to verify whether the labour supply specification based on varying effective wage rates (Models C, D) is preferred to the specification based on the unadjusted market wage rate for both household categories (Models A, B). These sets are not nested within each other. They differ in that the value of some of the right hand side variables are different for a subset of the observations. This can be treated as a difference in model specification when the equation is estimated over the full sample of individual members belonging to both categories.

Given the multiplicative form of the effective wage (w/θ) for Category 2 households, the version of Model C that allows the intercept (α term) to vary for Category 1 and 2 households would be indistinguishable from Model B, were it not for the fact that PNLY is computed in a different manner in Model C (and D). Otherwise, the β coefficient for the wage term for Category 2 households in Model

C would differ from the β coefficient for Category 2 households in Model B by the constant θ.[18] Hence it would not be possible to identify whether Equation 7.3B.2 or Equation 7.3C.2 had been estimated.

The fact that model identification depends critically on the values of PNLY being different in these two specification is a potential problem since the differences in NLY computed on the basis of w or w/θ are likely to be minor. This problem is further compounded if PNLY turns out to be correlated with the error term of Equation 7.3 so that it would be necessary to use an instrumented version of PNLY for which the differences between PNLY(w) and PNLY(w/θ) could be even smaller. The Wu-Hausman test (Hausman 1979) is used to verify whether the wage rates and the non-labour income can be treated as exogenous variables. [19]

7.5 Labour Supply Regression Results: Male Family Members

Table 7.3 summarizes the preliminary regression results for male labour supply using the observed market wage rate for the full sample of individuals in all three household categories - net sellers, net buyers and autarchic households with respect to their labour market exposure. The results are presented for Model A (with common parameters for all households) and two versions of Model B that allow for different intercepts and slope dummies for the wage and PNLY variable. Model B1 allows only for an intercept dummy while Model B2 allows for both intercept and own-wage slope dummies.[20] The regression models of Table 7.3 include other

[18] If PNLY and all other variables in Models B and C were to be the same then the coefficient β^1 in Eq. 7.5.B1 would be the same as β in Eq. 7.5C.1, and β^2 in Eq. 7.5B.2 would be equal to β/θ in Eq. 7.5C.2; hence the two models could not be distinguished.

[19] The Wu-Hausman test for exogeneity of regressors is based on a test statistic which measures the difference in the values of the estimated parameters of a model that results when the suspected endogenous variables are substituted for by their instrumented values. Under the null of exogeneity, the difference in the estimated parameters should be small (Hausman 1978)

[20] The standard errors reported in Table 7.3 are OLS errors with White's heteroskedasticity consistent adjustments. The two-step error correction is not required because all variables are based on the observed market wage rate.

variables whose estimated coefficients are not reported in that table. Most of these other variables are dummy variables for ethnic groups and regions.[21]

Although the parameter estimates in Model A and B1 appear reasonable (positive own-wage effects and negative income effects on labour supply), Model B2 which allows for the full set of slope and intercept dummies is clearly superior to Models A and B1. Since Models A B1 and B2 are nested within each other the preferred specification can be determined by verifying the significance of the extra dummy variables with appear in Models B1 and B2. The full results for Model B2 (given in Appendix Table 7A.2) show that five out of the six intercept and slope dummies are highly significant leading to significantly different wage and income effects for the three household categories. The model fit also improves greatly as one allows for additional dummy intercept and slope terms. The adjusted R^2 increases from 0.29 for Model A to 0.41 for Model B2.

In Model B2 the own-wage coefficients are of the expected sign (positive) and significantly different from zero for Category 1 and 2 households. The own-wage coefficient is not significantly different from zero for autarchic households. This last result is as predicted.[22] There is a similar difference in the income effects for the three household categories. The coefficients on the non-labour income variable are significantly negative for Category 1 and 2 which is consistent with leisure being a normal good. The income effect on labour supply for autarchic households however has a positive sign, although the coefficient is not significantly different from zero. This discrepancy in the income effect of autarchic households is not a theoretically expected result. It does, however, add to the inference based on the own-wage

---

[21] The regression results for some specifications with the full set of variables are given in the Appendix Tables for this Chapter.

[22] The result of an insignificant own-wage coefficient for individuals in autarchic households is based on a specification where the cross-wage effect for male labour supply is significantly negative, even for autarchic households. If the female wage rate variable is dropped from Model B2, the own-wage effects of male labour supply are still significantly positive for Category 1 and 2 households. For Category 3 (autarchic) households the own-wage coefficient becomes negative, but its value is insignificant. So whether the female wage cross effects are included or not, the inference from Model B2 is similar. The labour supply behavior of individuals in autarchic households is insensitive to the observed market wage rate, while there is a significant positive effect in other household categories.

Table 7.3 Male Labour Supply Regressions with common market wage rates

Data subset: Males, All Categories (1,2,3)	**Model A**		**Model B1**		**Model B2**	
Estimation: OLS (heterosked.consistent errors)	common parameters		intercept dummies only		intercept & slope dummies	
VARIABLE name	*Estimated Coefficient*	*t-ratio*	*Estimated Coefficient*	*t-ratio*	*Estimated Coefficient*	*t-ratio*
Male wage rate	2.777	1.53	5.505	3.50		
(hire out) Category 1 (sample N_1 = 1240)					6.301	3.26
(hire in) Category 2 (sample N_2 = 813)					9.869	5.05
(autarchic) Category 3 (sample N3 = 489)					0.599	0.38
Female wage rate	-4.002	-1.94	-4.034	-2.16	-4.529	-2.37
Non Labour Income (x 100)	-1.631	-10.26	-0.621	4.31		
Category 1					-3.307	-6.88
Category 2					-0.369	-2.52
Category 3					0.439	1.01
FAMILY SIZE	0.885	2.45	1.331	3.93	1.313	2.76
Number Male workers	-10.003	-9.66	-10.148	-10.56	-9.728	-8.56
Number Female workers	0.612	0.61	0.104	0.15	0.487	0.53
AGE	1.343	3.37	1.125	2.97	1.052	2.81
AGESQ	-0.025	-5.26	-0.021	-4.66	-0.021	-4.53
EDUCYR	-1.145	-2.29	-0.325	-0.07	-0.360	-0.79
HOUSEHOLD HEAD DUMMY	7.023	1.66	3.271	0.82	4.058	1.01
SEASON DUMMY	-4.194	-3.02	-4.166	-3.22	-4.162	-3.27
INTERCEPT	108.2	11.45				
Category 1			106.4	11.98	110.5	11.56
Category 2			80.8	9.04	61.3	6.24
Category 3			68.8	7.50	87.5	7.94
Adjusted R Square		0.29		0.39		0.41
Standard error of the estimate (SIGMA)		35.20		32.73		32.21
Breusch-Pagan HETEROSKESD. TEST	Chi. Sq. (df 19)	44.60 *	(df 21)	68.28 *	(df 25)	148.80
RESET (2) TEST	F (df 1, 2523)	64.09 *	(df 1, 2521)	5.09 *	(df 1, 2517)	6.18
Wu-Hausman Test for Exogeneity of						
Non-labour income	Chi. Sq.				(df 25)	18.60
Non labour income and wage rates	Chi. Sq.				(df 25)	27.22

* signifies the relevant test statistic is significant at the 5% significance level.

effects discussed above suggesting that the labour supply behaviour of the male workers in autarchic households appears to be quite distinct from that of individuals in households that report some labour selling or buying.

The non-labour income variable used in the regressions of Table7.3 is the actual value derived from valuing all family labour at the market wage rate. The Wu-Hausman test for the exogeneity of the non-labour income variable (as well as the wage rate) is reported in Table7.3.for Model B2 – the preferred model. The test statistic is not significant, indicating these variables are not correlated with the error term.[23] No specification bias results from using the computed values of per capita non-labour income variable; hence, instrumental variable techniques need not be used for the non-labour income variable computed on the basis of the observed market wage rates. This is a useful result that facilitates identification between Models B and C subsequently.

The coefficients on other regression variables do not differ greatly between the three models in Table 7.3. The observed relationships with these other variables are plausible. Individual labour supply increases with age but at a decreasing rate (implied by the negative coefficient on the age-squared term). The coefficient on years of education is negative in Model B2 but it is not significantly different from zero. The cropping season dummy is significant, with a slightly smaller labour supply intercept for the winter cropping cycle. The (male) household head supplies extra labour compared to other male family members. Controlling for family size, a larger number of available male family workers reduces the labour supply of each individual male worker. But a similar relationship does not occur with respect to the number of female workers in the household. This is an indication of work sharing within gender groups.

[23] The critical value of the Wu-Hausman test statistic (which is χ^2 with 25 degrees of freedom) is 37.65 at the 5% level of significance, and 34.38 at the 10% significance level. Since the computed values of the test statistic in Table 7.3 are less than these critical values, the null of exogeneity is not rejected.

In spite of the reasonable parameter estimates in Table 7.3, the RESET specification test indicates substantial model mis-specification for all three models that are based on the observed market wage rate. The mis-specification error is the largest for Model A which imposes common parameters for all household categories.

Table 7.4 provides a comparison of the regression results for alternative model specifications based on assigning different effective wage rates for Category 2 households. The sample of individuals in autarchic households is dropped for these regressions because the effective wage rate is not identified in terms of the θ parameter for autarchic households. [24] Model B in Table 7.4 is the same specification as Model B2 in Table 7.1 with the autarchic observations dropped. Models C and D use the w/θ effective wage for individuals in labor hiring households. All three specifications allow intercept and non-labour income slope dummies since these were significant effects in the preliminary results in Table 7.3. With this specification, the only difference between the three models presented in Table 7.4 relates to the definition of the effective wage rate and whether or not own-wage slope dummies are allowed.[25]

In comparing Models C and D, which are nested, the own-wage coefficients for labour selling and labour hiring households are not significantly different. Once the higher effective wage rate is allowed for Category 2 households, there is no significant difference in the own-wage slope parameter for individuals in Category 1 and 2 households.[26] Hence, Model D reduces to Model C.

---

[24] The autarchic households contain 489 records out of a total of 2542 person-season records in the full sample. Dropping the sample of individuals from autarchic households does not create a specification bias for the estimates reported in Table 7.4 since the labour supply coefficients are assumed to be different for individuals in the autarchic households.

[25] Table 7.4 does not report the estimate and standard errors for the actual intercept and slope dummy variables. These dummy variables have been added to the estimated base category parameters to compute the actual parameter value for each category. The dummy variable coefficients for several specifications are reported in the Appendix Tables where the full regression results are given.

[26] In Model D the dummy variable for the own-wage slope coefficient for Category 2 households is minus 0.129 with an adjusted standard error of 1.68. This value is not significantly different from zero.

Table 7.4 Male Labour Supply Regressions with alternative effective wage rates

Data subset: Males, Categories 1& 2 (excludes autarchic) **Estimation:** OLS (heterosked.consistent errors)	**Model B** common market wage intercept & slope dummies		**Model C** varying effective wage no wage slope dummy		**Model D** varying effective wage intercept & slope dummies	
VARIABLE name	*Estimated Coefficient*	*t-ratio*	*Estimated Coefficient*	*t-ratio*	*Estimated Coefficient*	*t-ratio*
Male wage rate			5.687	3.53		
(hire out) Category 1 (sample N_1 = 1240)	8.861	3.75			5.796	3.06
(hire in) Category 2 (sample N_2 = 813)	12.469	5.25			5.667	5.05
Female wage rate	-8.556	-3.46	-1.691	-1.05	-1.720	-1.06
Non Labour Income (x 100)						
Category 1	-3.435	-5.82	-3.125	-6.30	-3.217	-6.27
Category 2	-0.468	-3.06	-0.448	-2.98	-0.448	-2.97
FAMILY SIZE	0.770	1.89	1.044	2.76	1.043	2.76
Number Male workers	-10.289	-8.80	-9.433	-8.58	-9.428	-8.56
Number Female workers	1.622	1.47	1.896	1.87	1.897	1.87
AGE	1.658	3.74	1.398	3.32	1.398	3.32
AGESQ	-0.028	-5.32	-0.023	-4.59	-0.023	-4.59
EDUCYR	-0.325	-0.61	0.065	0.12	0.063	0.12
HOUSEHOLD HEAD DUMMY	7.023	1.66	4.040	0.90	4.029	0.90
SEASON DUMMY	-4.062	-2.75	-4.330	-3.11	-4.339	-3.11
INTERCEPT						
Category 1	110.1	9.72	94.1	10.11	94.0	9.80
Category 2	61.2	5.98	47.1	4.42	46.7	4.06
Adjusted R Square		0.32		0.39		0.39
Standard error of the estimate (SIGMA)		33.35		31.54		31.53
Breusch-Pagan HETEROSKESD. TEST	Chi. Sq. (df 22)	95.17 *	(df 21)	44.02 *	(df 22)	44.09 *
RESET (2) TEST	F (df 1, 2031)	2.17	(df 1, 2032)	2.03	(df 1, 2031)	2.04
Model selection Diagnostics						
Akaike Final Prediction Error		1125.3		1005.7		1006.7
Schwartz Criteria		1199.3		1071.8		1075.8

* signifies the relevant test statistic is significant at the 5% significance level.

In comparing Model C with Model B (which uses the observed market wage rate with varying slopes), Model C is the preferred specification on the basis of several diagnostic statistics - higher adjusted R square, smaller Akaike prediction error and also on the basis of the *J* test for non-nested models reported in Table 7.5.[27] In the preferred Model C, the Wu-Hausman test for exogeneity of the non-labour income variable based on the effective wage rates – PNLY(w/θ)- is insignificant again.[28]

Table7.5 Male Labour Supply: Non-nested *J* tests between Model B and Model C [29]

	Additional Variable	Estimated Coefficient	Stand. error	*t* statistic	Inference
Model B extended	predicted values from Model C	1.38	0.33	4.16	Reject Model B in favour of C
Model C extended	predicted values from Model B	0.56	0.42	1.35	Do not reject C in favour of B

Clearly Model C is the preferred specification over Model B.

Although Model C is preferred on statistical grounds, the comparison of the income and wage elasticities based on Model C and Model B given in Table 7.6 do not lead to very striking differences, with one exception. The cross wage elasticity with

[27] The standard procedure for carrying out the Davidson and MacKinnon (1981) *J* test for two competing non-nested models is to run an extended regression for each model that includes as an extra regressor the predicted values from the competing model. Model selection is based on the significance of the coefficient on the predicted value variables. This test does not always guarantee that one model will be preferred to the other (Maddala 1992: 515).

[28] The Wu-Hausman test statistic for the *joint* exogeneity of the non-labour income and effective wage rates in Model C was 24.22. This is less than the critical value of χ_2 with 21 degrees of freedom at the 5% and 10% significance levels. The test statistics for exogeneity tests of the non-labour income and the effective wage rates separately were also insignificant.

[29] The standard errors in Table 7.5 are based on OLS with White's heteroskedasticity adjustments only. They do not account for θ being a pre-estimated (as is done in Table 7.6). The two step error correction has little bearing on the model selection diagnostics such as the J tests because the adjustment in the standard error of specific parameters are very minor. Changes occur in the error for the real wage variable only.

respect to the female wage rate is insignificant in Model C while it is significantly negative in Model B. Table 7.6 also presents the elasticity values computed from a Model C specification that drops the female real wage variable from the male labour supply equation. This version of Model C restricts the cross-wage elasticity of male labour supply with respect to the female wage rate to be zero.

Table 7.6 Estimated Wage and Income Elasticities of Male Labour Supply

(standard errors below in italics)

Model B : common market wage and varying slope parameters

Household:		Elasticity with respect to: own wage	income	female wage	compensated own wage
(hire out)	Category 1	0.355	-0.078	-0.323	0.644
		0.095	*0.013*	*0.0934*	*0.109*
(hire in)	Category 2	0.675	-0.067	-0.438	0.671
		0.128	*0.0219*	*0.1206*	*0.128*
Model C	: varying effective wage; no wage slope dummies				
	Category 1	0.228	-0.070	-0.064 #	0.488
		0.065	*0.011*	*0.061*	*0.078*
	Category 2	0.410	-0.061	-0.113 #	0.404
		0.116	*0.203*	*0.101*	*0.116*
Model C	: varying effective wage; no wage slope dummies; **no cross-wage effect**				
	Category 1	0.181	-0.071		0.443
		0.057	*0.011*		0.074
	Category 2	0.326	-0.061		0.320
		0.102	*0.021*		*0.102*

Note: Computed at mean of data excluding individuals in autarchic households
# denotes not significantly different from zero.
Category 1 cases are individuals in households which are net sellers of labour
Category 2 cases are individuals in households which are net buyers of labour

The results for these two alternative specifications of Model C are similar with a slightly higher own-wage effect for the specification that allows for a negative cross-wage effect. [30] The own wage elasticities in Model B give slightly higher estimates than those based on either version of Model C. The estimates for the income elasticities are very similar in all model specification as well as for both categories of households. The computed income elasticity values are very low in all cases – less than 0.1 in absolute value.

Table 7.6 shows significant differences in the computed elasticities at the mean of the data for male workers in Category 1 and 2 households. The higher positive valued uncompensated own-wage elasticities for Category 2 are consistent with the theoretical prediction that a backward bending labour supply curve (negative own wage elasticities) is possible only for net sellers of labour (Category 1).[31]

The full results for Model C with and without the female wage variable are presented in Appendix Table 7A.3 with the 2 step estimator error correction to account for the fact that θ is a pre-estimated parameter with some error. In addition to the expected signs on the wage and income variables, the effects of the other variables are also reasonable. Holding family size constant, a higher number of male workers in the family reduces the workdays of an individual worker; but there is no corresponding effect from the number of female workers in the family. Work-days increases with age but at a decreasing rate. The diagnostic statistics indicate that in spite of the inflexible responses inherent in a linear labour supply equation, the RESET test does not detect functional form mis-specification in Model C (nor B).[32]

---

[30] Standard errors of the elasticity values noted in Table 7.6 for Model C are based on the 2 step adjusted covariance matrix accounting for the fact that that θ is a pre -estimated parameter. For the details of this procedure see Section 4.4 in Chapter IV.

[31] Strauss (1986:76)

[32] This result indicates that the source of the significant RESET test in Table 7.3 was the sample of individuals in autarchic households. Their labour supply behaviour is apparently not correctly captured by a linear specification, even when allowing for specific intercept and slope dummies with respect to the market wage rate and non-labour income variables. There is a strong possibility that individuals in autarchic households could be constrained on their off-farm labour supply behaviour.

In summary, the specification of a labour supply function for male family members that is consistent with the prior estimate of the higher efficiency of family labour in own farm production performs better than a model which assumes a common wage and constant parameters for all households.

7.6 Labour Supply Regression Results: Female Family Members

This section presents the results of the labour supply regression for female family members of the sample households. The estimation of the female labour supply regressions has an additional step to correct for the sample selection bias that may result from the large proportion of individual female family members who are reported to be economically inactive in the MPHBS employment data.

Out of the 1688 female family members enumerated in the sample households, almost 30% (481) individuals are economically inactive, according to the definitions adopted in the MPHBS.[33] This is a significant proportion. That sample selection bias may occur when the labour supply equations are estimated with the sample of only the economically active individuals is well known (Heckman 1979, 1980). To correct for this possibility, Heckman's two-step estimation procedure (Heckit) is used for the labour supply estimation for female family members.

In the first step of the Heckit procedure, a probit model for labour force participation is estimated using the full sample of economically active and inactive female individuals. In the second step, the labour supply (work-days) equation is estimated for the sample of economically active women only, but with an added explanatory variable (the inverse Mill's ratio) which controls for the possible correlation of the error terms between the participation and labour supply equations.[34]

[33] See Chapter V, Section 5.3(b) for the definitions for economically active persons adopted in the MPHBS data and for the categories of work identified.

[34] The source of the sample selection bias is not due to the lack of randomness in the reduced sample of the economically active individuals only. Rather, the problem occurs due to the correlation between the error terms in the model that determines labour market participation and the model that determines the amount of labour supplied. If there is no correlation between these two error terms, the inverse Mill's ratio variable is insignificant, the estimates based on OLS are consistent (Heckman 1979).

(a) probit labour force participation model

The maximum likelihood probit estimate of the female labour-force participation equation is given in Table 7.7(a). The binary dependent variable is defined so that those who are economically active are coded as 1, therefore a positive coefficient means an increase in the probability of labour market participation. Since the estimated probit coefficients do not directly give the marginal effects of a change in the explanatory variables, these are computed separately in Table 7.7(a) at the mean of the data, together with the implied elasticities.[35]

A proper implementation of the Heckit procedure requires that some of the explanatory variables in the probit equation be unrelated to the variables in the labour work-days equation (Heckman 1980). While a completely independent set of instruments is not available, the variable definitions used to denote non-labour income in the probit equation are different. In addition, since the Heckit two-step method already includes a special procedure for adjustments in the variance-covariance matrix for the parameters of the labour supply equation in the second step, the labour force participation equation is modeled to be independent of the θ parameter estimated in the production function. The wage rate variable used in Table 7.7(a) is the observed market wage. A dummy variable category for households that use hired labour is created. Similarly the non-labour income variable excludes farm profit so that an effective wage valuation of family labour need not be made at this step. A proxy for the imputed farm profit component of non-labour income is gross harvested area, with a dummy variable to control for households that are tenants. The returns to farm operation will be substantially less for tenant households than for owner-cultivators due to land rental payments.

The probit regression results indicate that a higher market wage rate, a higher non-labour income and a larger farm size increase the probability of labour market participation. The squared term of area harvested in significantly negative, implying the effect of increasing farm size will eventually be negative. Age has a similar

[35] The derivation of the marginal effects from the probit parameters follows Greene (1993:643-46).

Table 7.7(a) **Female Labour Force Participation Equation**
Data subset: **Female all categories, including inactive persons** (N = 1688)
Estimation method : Probit Maximum Likelihood

VARIABLE	*Estimated Coefficient*	*Asymptotic Standard Error*	*T-Ratio*	*Marginal Effects at means*	*Elasticity at means*
Female real wage rate	0.1154	0.055	2.11	*0.0360*	*0.1997*
Family size	0.0236	0.015	1.57	*0.0074*	*0.0862*
Number Male workers	0.0410	0.038	1.07	*0.0128*	*0.0355*
Number Female workers	-0.1639	0.039	-4.22	*-0.0512*	*-0.1734*
Age	0.1029	0.013	7.82	*0.0321*	*1.3783*
Age Squared	-0.0015	0.000	-8.47	*-0.0005*	*-0.7412*
Non-labour Household income (excludes imputed farm profit)	0.0022	0.001	2.35	*0.0007*	*0.0481*
FARM AREA harvested	0.4081	0.223	1.83	*0.0001*	*0.0667*
FARM AREA harvested squared	-0.0004	0.000	-3.99	*0.0000*	*-0.0582*
CONSTANT	-0.7949	0.347	-2.29	*-0.2481*	*-0.3273*
Dummy Variable Categories :					
Western Region	0.6405	0.080	7.96	*0.1999*	*0.0552*
Far-western Region	-0.7312	0.140	-5.23	*-0.2282*	*-0.0452*
Household Head	0.2909	0.183	1.59	*0.0908*	*0.0029*
Unmarried dummy	0.1651	0.203	0.81	*0.0515*	*0.0042*
Labor hire in dummy	-0.2377	0.363	-0.66	*-0.0742*	*-0.0371*
Presence of children aged 0-5	-0.1169	0.081	-1.44	*-0.0365*	*-0.0389*
Presence of children aged 6-9	0.0094	0.066	0.14	*0.0029*	*0.0022*
Tenant Household	0.3572	0.114	3.14	*0.1115*	*0.0155*

```
LOG-LIKELIHOOD FUNCTION =   -1369.5
LOG-LIKELIHOOD(0)   =       -1727.9
LIKELIHOOD RATIO TEST  =    716.806   WITH    29  D.F.

MADDALA R-SQUARE                          0.220
CRAGG-UHLER R-SQUARE                      0.315
MCFADDEN R-SQUARE                         0.207
     ADJUSTED FOR DEGREES OF FREED        0.199

           PREDICTION SUCCESS TABLE
                      ACTUAL
                  0              1

              0    199.           86.
 PREDICTED 1     282.          1121.

 ALL CASES       481           1207.

 NUMBER OF RIGHT PREDICTIONS =      1320
 PERCENTAGE OF RIGHT PREDICTIONS =   0.78
```

non-linear effect on labour market participation, with the age term being positive and the age squared term being negative. Other demographic variables also have plausible effects. The coefficients on family size and the number of male workers are positive, but not significant. A larger number of female workers in a household, however, reduces the probability of labour market participation for a particular individual. This implies unequal work sharing among female members in a household. The presence of children aged 0-5 in a household reduces the probability of female labour market participation, but there is no statistically significant effect of the presence of children aged 6 to 9. The dummy variable for tenant households is significantly positive. The coefficient on the dummy for hiring in labour is negative but insignificant. Several of the regional dummies are significant.

The model fit statistics indicate a reasonable fit for a binary dependant variable model. The Mcfadden R-sq. is 0.2. The prediction success table shows that while the overall percentage of right predictions is high (at 78%), the bulk of the right predictions are for individuals who are economically active. Among the 481 individuals who are not economically active, the number of right predictions is only about 41%. Most of the prediction errors result from not being able to correctly predict non-participation for specific individuals. Nevertheless, given that most of the explanatory variables used in the probit regression are household level variables, while the data show differences in the labour force participation among female members of the same household, the probit model performs reasonably well.

(b) labour supply (days worked) models

The labour supply regressions for female workers are also done only for individuals in Category 1 and Category 2 households. Observations from the autarchic households are dropped from the final regression because the effective wage rate is not observed for such individuals. Using the market wage rate for individual female workers in autarchic households leads to anomalous results, just as in the case of the labour supply regression for male household members.[36]

Another adjustment made in the female labor supply regressions is to drop the cross-wage term - the male wage rate variable. With both wage rates included in the regressions, neither becomes significant in most specifications. Since in the final specification of Model C for male workers, the female cross wage term was insignificant, it is also theoretically consistent to drop the male wage rate from the female labour supply regressions.[37]

The labour supply regression results for female family members are presented in summary form in Tables 7.7(b), (c) and (d) for model specifications B, C and D, respectively. Under each model specification there are two sets of parameter estimates: one based on OLS, ignoring the sample selection problem; the other is based on the second step of the Heckit procedure with the inverse Mill's ratio (IMR) as an additional variable.

Model B equates the opportunity cost of family labour to the observed market wage rate for female labourers, but allows for intercept and slope dummies for individuals in Category 1 and 2 households. The standard errors of the coefficients reported for Model B are White's heteroskedasticity consistent errors for the OLS estimates,

[36] In the version of Model B in Table 7.4, estimated over the sample of all female workers from the three households categories with appropriate intercept and slope dummies, the estimated own-wage coefficient for the labour supply of individuals in autarchic households is 0.525, with a standard error of 0.686. This estimate is not significantly different from zero, while the coefficients for the other two categories are significantly positive. Also, the RESET test statistic value of 18.985 indicated substantial mis-specification in Model B for female workers from all three household categories (as was the case in Model B for male workers).

[37] The precise theoretical restriction that is consistent with utility maximization is that the compensated cross-wage effects for male and female labour supply should be symmetric. (See the derivation in Kawaguchi (1994)). For this restriction to hold, it is not necessary that both of the uncompensated cross wage effects be zero, as is imposed by the specification above where the cross-wage variable is dropped from the labour supply equation for each gender. However, given that the income effects (the coefficients on the non-labour income variable) for both male and female individuals are approximately the same, the symmetry of the compensated wage effect depends crucially on the symmetry of the uncompensated effects. Hence, if one uncompensated cross-wage effect is zero, the other must also be close to zero in order to met the symmetry condition for the compensated cross-wage effect.

since heteroskedasticity is indicated in all model specifications. For the Heckit estimates of Model B, the reported standard errors are based on the standard adjustment required when the IMR is an additional variable because the IMR is also a generated regressor.[38] The Heckit standard errors do not explicitly control for hetero-skedasticity. Models C and D, however, use the effective wage rate variable based on the w/θ adjustment for Category 2 households. Hence, in Tables 7.7(c) and (d) an additional set of corrected standard errors is reported to account for the fact that θ is an estimated parameter from the production function equation.[39]

The parameter estimates of all specifications are very similar. The own-wage effect on female labour supply is significantly positive, and the non-labour income effect is significantly negative, as with the male labour supply results. The intercept dummy and the slope dummy for the non-labour income variable for Category 2 households is always significant. [40]

In all specifications the IMR variable is insignificant. The estimated correlation between the error terms in the labour force participation equation and the labour days equation is around 0.3. The Heckit model is still meaningful even though the coefficient on IMR is not significant. A likely reason for this insignificant coefficient is that the other regressors in the labour supply equation are similar to the regressors in the probit equation. This can lead to a high degree of multi-collinearity between the IMR variable and the other regressor in the labour supply equation, resulting in an insignificant coefficient for IMR. Because of this possibility the Heckit specification is preferred to the OLS.

---

[38] The procedure for deriving the Heckit corrected standard errors is given in the Shazam Manual (Shazam 1993:262-65).

[39] In these adjustments to the standard errors, each particular effect is taken to be additive. Hence the errors reported under the *hetcov* $+\theta$ *Adj.* columns for the OLS estimates in Table 7.7 are White's heteroskedasticity consistent errors (computed under the Hetcov option in Shazam) to which is added the positive definite matrix which results from the adjustment for the fact that θ is pre-estimated. (See Eq. 4.14 in Chapter IV). Similarly the Heckit corrected errors and the θ adjustment corrections are added together in the *heckit* $+\theta$ *Adj*. columns.

[40] The Wu-Hausman test for the exogeneity of the wage and income variables is not significant in the female workers sample either.

Table 7.7(b) Female Labour Supply Regression: Model B

Data subset: Categories 1& 2 (N = 1827) — common market wage

(excludes autarchic) — intercept & slope dummies

VARIABLE name	ols coef	*hetcov* se	*t ratio*	heckit coef	se	*t ratio*
Female wage rate						
(hire out) Category 1	4.194	1.828	*2.29*	4.742	1.794	*2.64*
(hire in) Category 2	8.061	1.670	*4.83*	8.016	1.670	*4.80*
Non Labour Income (x 100)						
Category 1	-2.289	0.511	*-4.48*	-2.280	0.469	*-4.87*
Category 2	-0.318	0.219	*-1.45*	-0.324	0.219	*-1.48*
FAMILY SIZE	2.810	1.123	*2.50*	2.825	1.162	*2.43*
Number Male workers	-7.265	1.385	*-5.25*	-7.035	1.505	*-4.67*
Number Female workers	-3.877	1.468	*-2.64*	-4.254	1.573	*-2.70*
No. of Children aged 0-5	-3.463	1.247	*-2.78*	-3.547	1.313	*-2.70*
AGE	1.316	0.393	*3.35*	1.680	0.546	*3.08*
AGESQ	-0.024	0.005	*-4.58*	-0.030	0.008	*-3.86*
EDUCYR	-1.732	1.020	*-1.70*	-2.855	1.751	*-1.63*
HOUSEHOLD HEAD DUMMY	13.891	4.748	*2.93*	14.844	4.602	*3.23*
SEASON DUMMY	-6.627	1.377	*-4.81*	-6.632	1.377	*-4.82*
Inverse Mills Ratio				**7.931**	**8.429**	***0.94***
INTERCEPT						
Category 1	45.0	10.4	*4.33*	35.7	14.3	*2.50*
Category 2	35.4	9.3	*3.81*	25.9	11.5	*2.25*
Adjusted R Square	0.300			0.300		
Standard error of the estimate (SIGMA)	29.68			29.68		
Log Likelihood	-8766.9			-8766.4		
Breusch-Pagan HETEROSKESD. TEST	166.7 *	Chi-Sq with 28 D.F.		168.58 *	Chi-Sq with 29 D.F.	
RESET (2) TEST	2.90	F with 1, 1799 D.F.		2.67	F with 1, 1798 D.F.	
Model selection Diagnostics						
Akaike Final Prediction Error	895.4			895.9		
Schwartz Criteria	980.2			983.8		
Correlation between participation and labour supply equations (ρ)				0.271		

Table 7.7(c) **Female Labour Supply Regression: Model C**

Data subset: **Categories 1& 2** (N =1827) varying effective wage

(excludes autarchic) no wage slope dummy

VARIABLE name	ols coef	hetcov se	hetcov + θ Adj. se	hetcov + θ Adj. t ratio	heckit coef	heckit se	heckit + θ Adj. se	heckit + θ Adj. t ratio
Female wage rate	5.120	1.200	1.274	*4.02*	5.374	1.257	1.370	*3.92*
Non Labour Income (x 100)								
Category 1	-2.356	0.502	0.503	*-4.69*	-2.326	0.465	0.466	*-4.99*
Category 2	-0.510	0.220	0.220	*-2.32*	-0.516			
FAMILY SIZE	2.852	1.122	1.122	*2.54*	2.885	1.160	1.160	*2.49*
Number Male workers	-7.322	1.386	1.386	*-5.28*	-7.067	1.497	1.499	*-4.71*
Number Female workers	-3.954	1.467	1.467	*-2.70*	-4.424	1.569	1.569	*-2.82*
No. of Children aged 0-5	-3.590	1.245	1.247	*-2.88*	-3.695	1.311	1.313	*-2.81*
AGE	1.333	0.393	0.393	*3.39*	1.764	0.528	0.529	*3.33*
AGESQ	-0.025	0.005	0.005	*-4.63*	-0.031	0.007	0.007	*-4.15*
EDUCYR	-1.700	1.014	1.015	*-1.67*	-3.037	1.700	1.708	*-1.78*
HOUSEHOLD HEAD DUMMY	13.775	4.777	4.777	*2.88*	14.909	4.576	4.577	*3.26*
SEASON DUMMY	-6.628	1.377	1.377	*-4.81*	-6.635	1.377	1.377	*-4.82*
Inverse Mills Ratio					**9.586**	**8.046**	**8.087**	***1.19***
INTERCEPT								
Category 1	41.6	8.9	9.1	*4.59*	32.0	12.1	12.4	*2.58*
Category 2	41.2	8.7	8.9	*4.63*	28.3	11.7	11.9	*2.38*

	ols		heckit	
Adjusted R Square	0.301		0.301	
Standard error of the estimate (SIGMA)	29.66		29.65	
Log Likelihood	-8766.1		-8765.4	
Breusch-Pagan HETEROSKESD. TEST	166.8 *	Chi-Sq with 27 D.F.	168.7 *	Chi-Sq with 28 D.F.
RESET (2) TEST	3.66	F with 1, 1800 D.F.	3.26	F with 1, 1799 D.F.
Model selection Diagnostics				
Akaike Final Prediction Error	893.7		894.0	
Schwartz Criteria	975.4		978.7	
Correlation between participation and labour supply equations (ρ)			0.325	

Table 7.7(d) Female Labour Supply Regression: Model D

Data subset: Categories 1& 2 (excludes autarchic) — varying effective wage, intercept & slope dummies

		ols	hetcov	hetcov + θ Adj.		heckit		heckit + θ Adj.	
		coef	*se*	*se*	*t ratio*	*coef*	*se*	*se*	*t ratio*
Female wage rate									
(hire out)	Category 1	4.062	1.832	1.833	*2.22*	4.622	1.791	1.792	*2.58*
(hire in)	Category 2	5.889	1.251	1.557	*3.78*	5.853	1.497	1.722	*3.40*
Non Labour Income (x 100)									
	Category 1	-2.306	0.510	0.510	*-4.52*	-2.296	0.468	0.468	*-4.91*
	Category 2	-0.507	0.219	0.223	*-2.27*	-0.513	0.253	0.253	*-2.03*
FAMILY SIZE		2.890	1.123	1.123	*2.57*	2.905	1.161	1.161	*2.50*
Number Male workers		-7.380	1.386	1.386	*-5.32*	-7.145	1.503	1.503	*-4.75*
Number Female workers		-3.989	1.469	1.469	*-2.72*	-4.376	1.571	1.571	*-2.79*
No. of Children aged 0-5		-3.599	1.245	1.247	*-2.89*	-3.685	1.311	1.312	*-2.81*
AGE		1.312	0.393	0.393	*3.34*	1.684	0.545	0.545	*3.09*
AGESQ		-0.024	0.005	0.005	*-4.57*	-0.030	0.008	0.008	*-3.87*
EDUCYR		-1.648	1.015	1.015	*-1.62*	-2.797	1.748	1.748	*-1.60*
HOUSEHOLD HEAD DUMMY		13.731	4.759	4.759	*2.89*	14.706	4.589	4.589	*3.20*
SEASON DUMMY		-6.629	1.377	1.377	*-4.81*	-6.634	1.377	1.377	*-4.82*
Inverse Mills Ratio						**8.121**	**8.421**	**8.422**	***0.96***
INTERCEPT									
	Category 1	46.0	10.4	10.4	*4.41*	36.4	14.3	14.3	*2.56*
	Category 2	37.6	12.0	12.0	*3.14*	27.9	12.0	12.0	*2.32*
Adjusted R Square			0.301			0.301			
Standard error of the estimate (SIGMA)			29.64			29.64			
Log Likelihood			-8765.7			-8765.2			
Breusch-Pagan HETEROSKESD. TEST			167.8 *	Chi-Sq with 28 D.F.		169.4 *	Chi-Sq with 29 D.F.		
RESET (2) TEST			3.42	F with 1, 1799 D.F.		3.18	F with 1, 1798 D.F.		
Model selection Diagnostics									
Akaike Final Prediction Error			894.2			894.8			
Schwartz Criteria			979.0			982.5			
Correlation between participation and labour supply equations (ρ)						0.276			

The tests for model selection below as well as the computed values of the various elasticities of female labour supply are based on the Heckit specification. Since the differences in the estimated parameters with and without the IMR variable are small, this choice of a preferred specification is not very consequential.

In terms of model selection, as before, Models C and D are nested but B is an independent model with a different wage variable.[41] The additional own-wage slope dummy coefficient for Category 2 households in Model D is insignificant in both the OLS and Heckit estimates[42]. Therefore Model D reduces to Model C with a common wage coefficient for individual households in Categories 1 and 2, but with the effective wage rate being w/θ for individuals in Category 2 households. As before, the model selection choice is between B (with a common market wage rate and varying wage coefficients) and Model C (with an adjusted effective wage rate and common coefficients) with the implied difference in the non-labour income variable as well.

In comparing Model C with Model B, the RESET test does not reject either specification. Model C has a higher adjusted R-square and a smaller residual variance than Model B that has one additional parameter. Because of this, Model C is preferred by the Akai prediction error and the Schwartz criteria (which are smaller for Model C). The non-nested *J* test also clearly shows that Model C is preferred to Model B, as indicated in Table 7.8. The full regression results for Model C for the female workers sample is given in Appendix Table 7A.4

---

[41] As in the case of the regressions for the male workers, all estimated versions of Models B, C and D for the female workers also allow for an intercept dummy and slope dummy for the non-labour income variable for Category 2 households. Therefore, the basis for choosing among the alternative model specifications for the female regressions will also depend solely on the definition of the effective wage rate variable and inclusion of the wage slope dummy variables.

[42] The estimated coefficient in Model D for the female wage rate slope dummy for Category 2 households is 1.82 with an standard error of 2.1 in the OLS specification, and 1.23 with a standard error of 2.2 in the Heckit specification. The p-values associated with the t-statistic for the test on these coefficients being significantly different from zero are 0.34 and 0.54, respectively. Therefore, the null of a common wage coefficient for all individuals, but based on a model with effective wage rates, is not rejected by a wide margin

Table7.8 Female Labour Supply: Non-nested *J* tests between Model B and Model C

	Additional Variable	Estimated Coefficient	Stand. error	*t* stastistic	Inference
Model B* extended	predicted values from Model C	6.91	1.02	6.77	Reject Model B in favour of C
Model C * extended	predicted values from Model B	-2.06	1.69	1.25	Do not reject C in favour of B

* Based on the Heckit specification for both models, using unadjusted standard errors

While Model C is the preferred specification, the resulting differences between the elasticities of labour supply between Models B and C are fairly minor (unlike the case for the male labour supply elasticities where there were some significant differences). The labour supply elasticities for female workers are given in Table 7.9 using the coefficients of the Heckit specification.[43] The relatively large difference observed in Model B for the two household categories on the uncompensated and compensated own-wage elasticities is reduced slightly with the effective wage specification in Model C. But in both specifications the difference in the elasticity for individuals in Category 1 and Category 2 households are statistically significant, even though both elasticities are derived from the same estimated parameter.[44] The higher own wage elasticity of labour for Category 2 in comparison to Category 1 households (which consist mainly of landless and small farm households) is the expected result (Rosenzweig 1980).

[43] The elasticities reported in Table 7.9 are conditional elasticities. They indicate only the effect on the labour supply of women who are already participating in the labour market. In the Heckit specification the effect of, for instance, a wage change has two dimensions: it increases the probability of participation among in-active women, and it increase the work-days of those who are already working (McDonald and Moffitt 1980). The first of these effects is ignored on the calculations of the elasticities in Table 7.9.

[44] The difference arises because the mean levels of the average days of work are quite different for individuals in Category 2 and Category 1 households. See Table 7.2.

In comparing Table 7.9 with the corresponding elasticities of male labour supply given in Table 7.6, the own-wage elasticities under the preferred Model C are higher for female workers than for males. Similarly the income elasticity of labour supply is also slightly higher for female workers. A higher wage and income elasticity for female labour supply compared to male workers is an expected result commonly found in both developed and developing country settings (Killingsworth and Heckman 1986, Bardhan 1979).

Table 7.9 Estimated Wage and Income Elasticities of Female Labour Supply

(standard errors below in italics)

Model B: common market wage, varying slope parameters; no cross-wage effect

Household Category:		Elasticity with respect to: own wage	income	compensated own wage
(hire out)	Category 1	0.280	-0.0823	0.402
		0.098	*0.018*	*0.120*
(hire in)	Category 2	0.665	-0.0773	0.662
		0.149	*0.044*	*0.139*

Model C: varying effective wage; no wage slope dummies; no cross-wage effect

		own wage	income	compensated own wage
	Category 1	0.317	-0.0851	0.558
		0.075	*0.018*	*0.141*
	Category 2	0.594	-0.1064	0.431
		0.104	*0.043*	*0.082*

Note:computed at the mean of the data excluding individuals in autarchic households.

7.7 Summary

The labour supply regression results in this Chapter represent the second step in the sequential estimation strategy of a farm household model that allows for labour heterogeneity. This step had two objectives. First, to generate a complete set of the labour supply elasticities to describe, in aggregates form, the preferences of the farm households.[45] Second, to provide additional evidence for the labour heterogeneity results of Chapter VI by verifying whether the observed labour supply behaviour was theoretically consistent with the finding of a higher productivity of family labour.

An effective labour composite nest of the linear form ($Le = F + \theta H$) implies a higher effective wage is applicable for family labour in own-farm work as a substitute for hired labour. The underlying farm household model becomes recursive since the wage gap is given by a constant (θ), and the effective wage rate becomes exogenous to the household's production and labour supply decisions. A two step estimation strategy, however, is still required because some of the regressors in the labour supply equation are based on the θ parameter estimated in the production function.

An independent verification of the production function estimation result that $\theta < 1$ can be made by comparing alternative labour supply model specifications. For both the male and female workers sample, the model allowing for labour heterogeneity (Model C) is always preferred to the model based on homogeneous labour (Model B) with a common market wage rate for all households. This result is verified through several diagnostic statistics on model specification, including the *J* test for non-nested models which gives a clear verdict in favour of Model C.

In the specifications based on the effective wage rate, the wage slope dummy variable for labour hiring households is not significant (in Model D). This means the effective wage rates based on the θ parameter estimated from the production function are correctly defined. If the true effective wage rates for family labour in households using hired labour were substantially different from w/θ, the wage slope dummy variable in Model D would have been significant. This is additional evidence for the value of θ estimated in the production function being consistent with the labour supply behaviour of farm household members.

---

[45] These labour supply estimates in effect represent the consumer equilibrium of the farm household expressed implicitly in a demand system with a composite consumption good and leisure. Hausman (1981) has derived the direct and indirect utility functions underlying a linear labour supply function. Stern (1986) also discusses the utility implications of a linear labour supply function.

Another implication of $\theta < 1$ is that households should not engage simultaneously in hiring in and hiring out of labour. This prediction is strongly supported. Only about 3% of the sample household report simultaneous labour hiring and hiring out, even though the incidence of labour hiring in is not limited to big farmers.

Another important finding is that the estimated values for the wage and income elasticities for labour supply differ between the specifications based on the effective wage variable and those based on the market wage rate. The differences are more striking for the elasticities of male labour supply.

In addition to model selection tests, the overall regression results for the model based on effective wage rates are very reasonable. The estimated parameters indicate significantly positive own-wage effects and negative income effects; and female labour supply is more elastic than male labour supply with respect to both wage and income. There are also substantial differences in the wage and income elasticities of labour supply for individuals who are in households that are net buyers and net sellers of labour. These are in the theoretically predicted direction. The labour supply regression equations for autarchic households using the observed market wage rate do not give meaningful results, which is also a theoretically expected outcome.

In conclusion, a simple linear labour supply model with adjusted wage rates provides a reasonable description of the labour supply behaviour of farm-household members in the *tarai* region of Nepal.[46] The regression results for both male and female family members is consistent with the prior estimate of a higher efficiency of family labour in own farm production. These results provide independent corroboration for the labour heterogeneity indicated in the production function estimation.

---

[46] The linear labour supply equation specifications used in this Chapter are very simple. The prime interest was to find corroborating evidence for labour heterogeneity. Other interesting questions in the labour supply behavior of farm households - i.e., a more flexible response of labour supply to wage and income changes, intra-family labour allocation rules, female labour supply decisions being conditional on male labour supply, allowing for individuals to be constrained in their off-farm work-days, household fixed effects, etc. - were not addressed. These issues and other more elaborate analytical structures and estimation procedures for modeling the labour supply behaviour of farm households, together with labour heterogeneity, would be a fruitful area for further research.

APPENDIX 7.1

Complete Labour Supply Regression Results for Selected Model Specifications

This appendix presents the full set of parameter estimates and diagnostic statistics for some selected model specifications for which only summary regression results were presented in the main body of Chapter VII.

This appendix contains four tables that are as follows:

Appendix Table 7A.1 gives the definition of the variable names used in the complete regression results.

Appendix Table 7A.2 presents the full results for Model B2 of Table 7.3 in the main text of Chapter VII. This Model B2 is estimated for the sample of male workers from all three household categories - labour selling, labour hiring and autarchic, using the observed market wage rates for hired labour to represent the opportunity cost of family labour for all individuals.

Appendix Table 7A.3 presents the complete results for Model C of Table 7.4 in the main text of Chapter VII. This Model C is estimated for the sample of male workers from only the labour selling and labour hiring households, using the effective wage rates based on θ. Two versions of Model C are estimated, with and without the female cross-wage variable.

Appendix Table 7A.4 presents the complete results for Model C estimated for the sample of female family workers from the labour selling and labour hiring households, using the effective wage rates based on θ. This table is the full results that goes with Model C of Table 7.79(c) in the main text of Chapter VII.

Appendix Table7A.1

VARIABLE DEFINITIONS for Appendix Regression Tables

AGE		age of individual
AGESQ		age squared
CONS		CONSTANT
DISDM21		dummy variable for western region districts
DISDM23		dummy variable for far western region districts
DMRLHH3		dummy for own children of houeshold head
DMRLHH4		dummy for parents of household head
DMRLHH6		dummy for siblings of household head
DMRLHHI		dummy for household head
EDUCYR		
ETHOHDM		dummy for low caste status of Hill region origin
ETHTLDM		dummy for low caste status of Terai region origin
FMNOFLF2		number of economically active female adults aged 15 -59
FMNOMLF2		number of economically active male adults aged 15 -59
FMSIZE		
HHAUTDM		same as LS3
IMR		inverse Mill's ratio variable
LHIUSEDM		same as LS2
LS2		dummy variable for Category 2 households (net buyers of labour)
LS2XPFNT		LS2 interacted with PFNLTR42
LS2xPFNY		LS2 interacted with PFNLYR42
LS2xRWGM		LS2 interacted with RLWGRTM2
LS3		dummy variable for Category 3 households (autarchic)
LS3xPFNY		L:S3 interacted with PFNLYR42
LS3xRWGM		LS3 interacted with RLWGRTM2
LSI		dummy variable for Category I households (net sellers of labour)
NOC05		numer of children aged 0 to 5 in household
NOC69		numer of children aged 6 to 9 in household
PFNLYR42		real per capita household non labour income, including net farm profit computed by deducting value of own family labour at market wage rates
PFNLTR42		real per capita household non labour income, including net farm profit computed by deducting value of own family labour at effective wage rates given by RLWGF2TH and RLWGM2TH
PHASEDM		dummy for seasonal phase of survey data (2 phases)
RLWGF2TH	=	RLWGRTF2 for LS I category housheolds
	=	(RLWGRTF2/θ) for LS2 category housheolds
RLWGM2TH	=	RLWGRTM2 for LS I category housheolds
	=	(RLWGRTM2/θ) for LS2 category housheolds
RLWGRTF2		Daily wage rate for female hired labour, deflated by price of paddy
RLWGRTM2		Daily wage rate for male hired labour, deflated by price of paddy
UNMARRDM		dummy for unmarried person
Dependent variable	=	total labour days of work on own farm or in hired labour market reported on a monthly basis for each of two seasonal cropping cycles

Appendix Table 7A.2
Labour supply Regressions
Data subset: Males All categories (LS1, IS2, LS3)
Model B2 common market wage, varying slopes and intercepts
Estimation method: OLS with White's heteroskedasticity consistent errors

	VARIABLE NAME	*ESTIMATED COEFFICIENT*	*STANDARD ERROR*	*T-RATIO 2517 DF*	*P-VALUE*
	RLWGRTM2	6.301	1.965	3.21	0.001
	RLWGRTF2	-4.529	1.910	-2.37	0.018
	FMSIZE	1.31	0.328	4.00	0.000
	PFNLYR42 (x100)	-3.307	0.481	-6.88	0.000
	FMNOMLF2	-9.782	0.949	-10.31	0.000
	FMNOFLF2	0.488	0.921	0.53	0.597
	AGE	1.052	0.375	2.81	0.005
	AGESQ	-0.021	0.005	-4.53	0.000
	DISDM21	9.00	1.698	5.30	0.000
	DISDM23	-29.55	3.233	-9.14	0.000
	PHASEDM	-4.16	1.271	-3.27	0.001
	DMRLHHI	4.06	4.030	1.01	0.314
	DMRLHH3	1.70	3.523	0.48	0.630
	DMRLHH4	4.50	6.455	0.70	0.486
	DMRLHH6	-3.064	3.935	-0.78	0.436
	EDUCYR	-0.36	0.455	-0.79	0.428
	UNMARRDM	-11.7	2.859	-4.08	0.000
	ETHOHDM	4.47	2.483	1.80	0.072
	ETHTLDM	4.87	1.949	2.50	0.012
Own wage slope dummies					
categ 2	LS2xRWGM	3.568	1.839	1.94	0.052
categ 3	LS3xRWGM	-5.702	1.651	-3.46	0.001
Income slope dummies (x100)					
categ 2	LS2xPFNY	2.938	0.5	5.88	0.000
categ 3	LS3xPFNY	3.746	0.65	5.77	0.000
	CONS	110.5	9.561	11.56	0.000
Intercept dummies					
categ. 2	LHIUSEDM	-50.6	7.986	-6.34	0.000
categ. 3	HHAUTD	-22.9	8.206	-2.80	0.005

R-SQUARE = 0.4109 R-SQUARE ADJUSTED 0.4051

STANDARD ERROR OF THE ESTIMATE-SIGMA 32.216
GOODNESS OF FIT TEST FOR NORMALITY OF RESIDUALS - 60 GROUPS
CHI-SQUARE = 40.3648 WITH 32 DEGREES OF FREEDOM

HETEROSKEDASTICITY TESTS
E**2 ON YHAT: CHI-SQUARE 4.360 WITH I D.F.
E**2 ON X (B-P-G) TEST: CHI-SQUARE = 148.808 WITH 25 D.F.

RAMSEY RESET SPECIFICATION TESTS USING POWERS OF YHAT
RESET(2)= 6.1795 -F WITH DF1= 1 AND DF2=2517

Appendix Table 7A.3
Labour Supply Regressions

Data subset:	**Males, Categories 1& 2** (excludes autarchic) (N = 2053)	
Model	**C**	varying effective wage, no wage slope dummies
Estimation method		OLS with White's heteroskedasticity consistent errors and the 2 step error correction for estimated θ

Version	1			2			
	with female wage			without female wage			
VARIABLES	*Coeff.*	*Stand. Error*	*T-ratio*	*Coeff.*	*Stand. Error*	*T-ratio*	*p-Value*
RLWGM2TH	5.687	1.614	3.52	4.521	1.417	3.19	*0.001*
RLWGF2TH	-1.691	1.604	-1.05	--	--	--	--
FMSIZE	1.044	0.378	2.76	1.066	0.462	2.31	*0.021*
PFNLTR42 x 100	-3.125	0.496	-6.30	-3.143	0.497	-6.32	*0.000*
FMNOMLF2	-9.433	1.099	-8.58	-9.476	1.136	-8.34	*0.000*
FMNOFLF2	1.899	1.014	1.87	1.870	1.037	1.80	*0.072*
AGE	1.399	0.421	3.33	1.391	0.421	3.31	*0.001*
AGESQ	-0.023	0.005	-4.59	-0.023	0.005	-4.57	*0.000*
DISDM21	8.325	1.792	4.65	8.029	1.894	4.24	*0.000*
DISDM23	-25.426	4.086	-6.22	-24.849	4.192	-5.93	*0.000*
PHASEDM	-4.334	1.393	-3.11	-4.335	1.394	-3.11	*0.002*
DMRLHH1	4.033	4.463	0.90	4.370	4.448	0.98	*0.326*
DMRLHH3	1.274	4.120	0.31	1.530	4.122	0.37	*0.711*
DMRLHH4	9.086	7.072	1.29	9.270	7.097	1.31	*0.192*
DMRLHH6	-4.774	4.526	-1.06	-4.274	4.491	-0.95	*0.341*
EDUCYR	0.065	0.500	0.13	0.079	0.502	0.16	*0.875*
UNMARRDM	-7.332	3.337	-2.20	-7.264	3.350	-2.17	*0.030*
ETHOHDM	3.065	3.119	0.98	3.104	3.118	1.00	*0.320*
ETHTLDM	4.425	2.110	2.10	4.351	2.143	2.03	*0.042*
Catgory 2 Income slope dummy							
LS2XPFNT x 100	2.676	0.511	5.24	2.696	0.514	5.25	*0.000*
CONS	94.1	9.810	9.59	92.400	10.630	8.69	*0.000*
Category 2 Intercept dummy							
LHIUSEDM	-47.0	5.310	-8.85	-47.300	8.155	-5.80	*0.000*
Adjusted R Square			0.39			0.39	
Standard error of the estimate (Sigma)			31.53			31.53	
Breusch-Pagan Heterosked.Test			44.03 reject null (χ^2 with Df = 22)			35.03 reject null (χ^2 with Df = 21)	
RESET (2) Test	(F with Df=1 & 2031		2.03		F (1, 203	2.50	
Model selection Diagnostics							
Akaike Final Prediction Error			1005.7			1005.2	

See Appendix Table 7A.1 for the variable definitions.

Appendix Table 7A.4
Labour supply Regressions
Data subset: **Female, Categories 1 and 2** (N = 1827)
Model **C** varying effective wage, common own-wage slope
Estimation method : Heckit

VARIABLES	Estimated Coefficient	Heckit Standard Error	Heckit T-Ratio 1798 Df	Heckit + θ Adj Standard Error	Heckit + θ Adj T-Ratio 1798 Df
RLWGF2TH	5.3737	1.2570	4.28	1.3700	3.92
FMSIZE	2.8845	1.1600	2.49	1.1600	2.49
PFNLTR42 (x 100)	-2.3258	0.4650	-5.00	0.4658	-4.99
FMNOMLF2	-7.0669	1.4970	-4.72	1.4990	-4.71
FMNOFLF2	-4.4239	1.5690	-2.82	1.5690	-2.82
AGE	1.7643	0.5275	3.35	0.5292	3.33
AGESQ	-0.0308	0.0074	-4.17	0.0074	-4.15
DISDM21	15.1070	2.6550	5.69	2.6600	5.68
DISDM23	-43.1060	4.1660	-10.35	4.1950	-10.27
PHASEDM	-6.6350	1.3770	-4.82	1.3770	-4.82
DMRLHH1	14.9090	4.5760	3.26	4.5770	3.26
DMRLHH3	8.6481	4.9620	1.74	4.9640	1.74
DMRLHH4	6.9594	3.9040	1.78	3.9050	1.78
DMRLHH6	21.0860	7.1280	2.96	7.1430	2.95
EDUCYR	-3.0366	1.7000	-1.79	1.7080	-1.78
UNMARRDM	-3.2926	4.9530	-0.66	4.9560	-0.66
LS2XPFNT (x 100)	0.0181	0.0052	3.49	0.0052	3.48
ETHOHDM	7.6905	3.2150	2.39	3.2150	2.39
ETHTLDM	3.4114	2.5610	1.33	2.5640	1.33
NOC05	-3.6948	1.3110	-2.82	1.3130	-2.82
NOC69	0.0376	1.2930	0.03	1.2930	0.03
ETHTHDM	-0.0365	2.4620	-0.01	2.4730	-0.01
ETHRUDM	16.6350	3.4730	4.79	3.4730	4.79
ETHUNDM	-2.7452	2.6810	-1.02	2.6820	-1.02
LHIUSEDM	-17.9	6.6910	-2.68	6.7820	-2.64
IMR	0.0959	0.0805	1.19	0.0809	1.19
CONS	46.4	12.110	3.83	12.420	3.74

Adjusted R Square	0.3013
Standard error of the estimate (SIGMA)	29.65
Log Likelihood	-8765.4
Breusch-Pagan HETEROSKESD. TEST (Ch.Sq. with Df=28)	166.8
RESET (2) TEST (F with Df= 1, 1798)	3.56
Model selection Diagnostics	
Akaike Final Prediction Error	894.0
Schwartz Criteria	978.7

CHAPTER VIII

SUMMARY AND CONCLUSIONS

8.1 The Research Question and Motivation

The main research question addressed in this thesis is whether it is important to distinguish between family and hired labour as production inputs in the traditional peasant agricultural production systems of the southern plain (*tarai*) region of Nepal. The prime motivation behind this question is to test the validity of the conventional specification of the farm household model that treats family and hired labour as homogeneous inputs, allowing the production and consumption decisions of farm households to be modeled recursively. Another motivation is to explain the variation in labour input across farm size so commonly observed in traditional agriculture.

Farm household models that reflect the integrated complex behavioral responses of a joint producer and consumer agent are an important analytical tool for policy makers in developing countries. A recursive farm household model, in which production and consumption/labour supply decisions need not be modeled jointly, offers considerable empirical estimation advantages. Econometric estimation of non-recursive models is analytically cumbersome and the behavioural parameters of interest often involve non-linearities in the estimating equation even if the underlying model is linear in both the production and labour supply components.[1] Hence, the recursive feature has great practical advantages and increases the popularity of farm household models for policy applications. Nevertheless, it is important to test that the empirical basis for the recursive structure is indeed well founded in a wide range of country settings. While there are several other grounds under which the separable property of farm household models breaks down, the main concern is the completeness of rural labour markets and competitive wage determination, and whether hired and family labour inputs are homogeneous inputs in farm production.

---

[1] This potential problem has been highlighted by Jacoby (1993: 908 footnote 5).

When family and hired labour must be treated as heterogeneous inputs the recursive structure of production and consumption choices usually breaks down. There will be a separate supply and demand equilibrium for family labour in which the household is no longer a price taker. The labour supply of the household is affected by its production decisions, and similarly its factor demands are affected by the household's consumption/leisure preferences.[2]

8.2 Methodology

The estimation methodology follows a two step procedure. In the first step an aggregate farm production function is estimated through which the heterogeneity of family and hired labour is tested. Several alternative specifications that allow for imperfect substitution between family and hired labour and varying marginal products are estimated. Heterogeneity tests of the two types of labour are carried out through standard tests of statistical significance of the parametric restrictions that lead to a model with homogeneous labour inputs. The estimated parameters of the preferred production function are also be used to derive the relevant factor demand elasticities that define the production side of the behavioural response of the farm household.

In the second step, a structural labour supply model is specified that is consistent with the type of labour heterogeneity detected in the production function estimation. This typically means the variables used in the labour supply regressions will be derived from parameters estimated in the production function. The main variable of interest derived through this process is the appropriate "shadow" wage rate that reflects the true cost of family labour at the equilibrium labour supply position of different types of households. A key question is how this shadow wage rate is related to the observed market wage rates for family and hired labour. The answer usually varies for different households based depending on whether they are net

[2] As noted in Chapter III, an intuitive way to understand the recursive property is that it necessarily holds as long as there are markets for all commodities and inputs, and the household is a price taker in all these markets, and the household owned or produced inputs and commodities are perfect substitutes for the market purchased versions.

buyers or sellers on the hired labour market. The labour supply equations are then estimated with these appropriate shadow or effective wage rates and the other variables that may also depend on the computed effective wage rates. The parameter estimates from these labour supply regressions together with the correctly specified variables then describe the labour supply behaviour of the sample household in a manner consistent with utility maximization and with the specific type of labour heterogeneity modeled in the production function.

The income and wage elasticities derived from the labour supply regression functions complete the set of parameters of the farm household model. Accurate estimates of the parameters of the labour supply function are of interest in themselves. However, these labour supply regressions can also be used to verify whether the labour supply model specification, based on the assumption of labour heterogeneity is superior to the conventional specification, based on the assumption of homogeneous labour. In the latter specification the observed market wage rate for hired labour would be an appropriate measure also of the opportunity cost of family labour. Hence, the comparative statistical performance of alternative labour supply model specifications that allow for a common wage (consistent with family and hired labour being homogeneous inputs) and varying effective wages (consistent with labour heterogeneity), can independently corroborate the result that family and hired labour are not homogeneous inputs in farm production.

The two-step estimation procedure adopted in this thesis is based on the strategy proposed by Jacoby (1993) where he implements this method for estimating a fully non-recursive farm household model in which family and hired labour are treated as completely separate inputs. The same sequential approach can be used to estimate a farm household model structure by estimating, in the first step, a farm production function in which is embedded a test for the heterogeneity of family and hired labour. Then in the second step the labour supply function is specified and estimated in a theoretically consistent manner, if heterogeneity is indicated.

Herein lies the methodological novelty of this thesis. On the one hand, while the conventional approach to estimating farm production functions is simply to aggregate family and hired labour into a homogeneous total labour input, there are many exceptions which estimate production functions by treating family and hired labour as completely separate inputs, implying but not formally testing the extent to which they are imperfect substitutes. On the other hand, in labour supply estimations the conventional approach has been to assume that the observed market wage rate is the appropriate opportunity cost of labour for all individuals. But a relatively small number of studies have used alternative derivations of a "shadow" wage rate variable to model the on-farm component of the labour supply behavior of farm household members. The two parts of these exceptional treatments have not been combined directly in a theoretically consistent manner - i.e., to relate the shadow wage rates of family labour to the observed market wage rates for hired labour based on the extent of the efficiency differences between hired and family labour detected in the production function estimates. That was the task carried out in this study.

8.3 Data Source

The data utilized for the empirical component of this thesis come from a large nationally representative household budget survey conducted by Nepal Rastra Bank. The data collected in this survey has been recognized to be of a very high quality and comprehensive in its treatment of farm income, inputs and outputs (World Bank 1992). The actual household sample used in this study is a subset of the national survey data limited to about 1,000 rural households (of which about 700 have operational land holdings) from the southern plain (*tarai*) region of Nepal.

In the *tarai* region there is a greater inequality of land ownership and higher incidence of hired labour use than in the farms of the northern hill and mountainous regions of Nepal. The primarily subsistence farm households in the northern regions engage in a whole range of other ancillary activities apart from on-farm crop production. It is difficult to correctly specify the inputs and outputs for these other activities. As a result, the question of heterogeneity between family and hired labour in Nepalese agriculture is better addressed in the context of the *tarai* region.

The production function estimates are based on the aggregate annual crop output for the land-operating households. The labour supply regression are based on the seasonal work days reported by all economically active adults aged 15-64 in the full sample of households (including the landless) whose main occupation was reported as an agricultural worker or own-farm operator. This led to about 2500 individual-season labour supply records for the male family workers and 1830 person-season-records for the female workers in the final labour supply regressions.

8.4 Summary of Analytical Results

There are three separate requirements that must be fulfilled for two production inputs, X_1 and X_2, if they are to be homogeneous:

(a) X_1 and X_2 be separable[3] from other inputs in the production function

(b) the elasticity of substitution between X_1 and X_2 is very large

(c) the ratio of the marginal products of X_1 and X_2 is equal to one (at all levels of applications of X_1 and X_2). In other words, a one unit increase in the application of X_1 has the same effect on output, *ceterus paribus*, as a one unit increase in X_2 has.

The estimation results with respect to family and hired labour reported in this thesis do not reject conditions *a* and *b;* but *c* is clearly rejected.

The tests for the heterogeneity between family and hired labour were carried out in Chapter VI, using a translog production function specification. Two different sets of estimations were carried out. The first set was restricted to a sub-set of the sample

[3] For a production process utilizing *n* inputs, *separability* of inputs X_1 and X_2 from the other inputs implies that the marginal rates of substitution between X_1 and X_2 are independent of the levels of the other *n*-2 factors. This implies that the ratio of the marginal products of X_1 and X_2 is invariant to the level of the other inputs. This notion of "separability" is, of course, a completely different concept than the "separability" or recursive property of farm household models.

households that reported the use of both family and hired labour in crop production. For this sample sub-set the translog production function was estimated with family and hired labour as two distinct inputs in addition to three other inputs (land, bullock power, and material inputs) which were interacted with the two labour inputs.[4]

The results from the first set of estimations clearly indicated that family and hired labour inputs were weakly separable from the three other inputs. Stronger forms of separability, including the Cobb-Douglas restriction of the translog production function, were rejected. The weak separability result means that a unit change in the application of the latter three inputs leaves unchanged the ratio of the marginal products of family and hired labour. Another key finding was that the estimated value of the marginal product of family labour was statistically higher than the marginal product of hired labour, computed at the geometric mean of the data; and the marginal product of hired labour was close to the observed market wage rates.

The weak separability of the labour inputs in the production function implies that the two types of labour can be consistently aggregated into a composite labour input. In the second set of estimations, several alternative functional forms for the composite labour aggregator function were specified, and the translog function with a nested aggregate labour input was estimated utilizing the entire sample of land cultivating households, including those that did not use any hired labour inputs.

The empirical results showed that the preferred labour input aggregator function was a linear composite given by $Le = F + \theta H$, where Le is effective or composite labour, F is family labour days and H is hired labour days. The estimated value of θ was 0.75, and it was shown to be significantly less than one. This preferred form implies that, although family and hired labour are perfect substitutes in farm production, they are not equally productive. When both are measured conventionally in time units, the application of an extra unit of family labour has a larger effect on

[4] Each of the family and hired labour input sub-totals is an aggregation of male and female labour within in each category. This aggregation is done using the ratio of the reported village-level female wage rate to the male wage rate for hired labour to convert female labour days into equivalent male labour days. The mean value of this ratio was 0.85 in the sample data.

output than an extra unit of hired labour. When hired and family labour are measured in *effective* units, one unit of hired labour substitutes for 0.75 units of family labour.

The higher productivity of family labour inputs can be readily justified on the grounds that the *effort* applied per unit of time is likely to be lower for hired than for family labour when supervision of hired labour is costly (Feder 1985). Another reason could be that family labour acquires some farm-specific experience, which leads to higher labour productivity (Rosenzweig and Wolpin 1985). However, the estimation work of this study, was not designed to discriminate between alternative explanations for the lower productivity of hired labour. The detailed information required to do so was not available in the survey data used.

A key implication of the Le = F + θ H labour aggregator function is that the difference in productivity between hired and family labour is constant and unaffected by the levels of other inputs. The analytical structure of the farm household model with heterogeneous labour inputs derived in Chapter III showed that under such conditions the farm household model is still recursive. The only difference with the conventional model with homogeneous wages is that the effective wage rates for family labour will differ according to the labour market exposure of the household in the hired labour market.

For a landless or a small farm household which at the margin supplies labour on the hired labour market, the effective wage rate which determines its total labour supply equilibrium will be the observed market wage rate, w, for hired labour. For big farm households that are net buyers of hired labour, the effective wage rate they face for the supply of their own family labour is w/θ, since at the margin one unit of family labour can substitute for 1/θ units of hired labour.

The result that θ is less than one in the production function estimation implies it would be irrational for the sample households simultaneously to hire in and hire out labour. This would be inconsistent with the higher productivity of family labour. There would be efficiency gains from transferring the hired-out labour into own farm cultivation, substituting for the hired-in labour. This prediction is borne out by the sample data. Only about 3% of the approximately 700 land-operating households report hiring in labour as well as some member of the household working on the hired labour market. Given that the survey data refers to the entire annual cropping cycle, and given the very time specific nature of agricultural operations and the strict gender-related division of labour, this is a striking result which is consistent with the efficiency difference between family and hired labour.

The estimation results for the labour supply component of the farm household model, reflecting the θ efficiency difference between family and hired labour, was presented in Chapter VII. The labour supply variable includes all reported workdays on own farm cultivation as well as hired labour market work over each of two seasonal survey rounds. The labour supply regression equations were estimated separately for individual male and female family members without imposing any cross restrictions. A probit sample selection correction was made for the labour supply regression equation for female workers to take account of the large proportion of women who report to be economically inactive.

The labour supply regression results with the θ adjusted effective wage rates (Model C) are very reasonable for both the male and female regressions. The own wage elasticities of labour supply are positive (0.2 to 0.4 for male workers, and 0.3 to 0.6 for female workers) and within the range of estimates obtained by others for farm households in developing countries. Labour supply decreases with the level of non-labour income (implying leisure is a normal good); but these elasticities are rather small (-0.07 to -0.1), which is again consistent with previous studies. The differences in the estimated elasticities between the two genders and between small (labour hiring-out) and big farm (labour hiring-in households) are also as expected. The absolute values of the wage and income elasticities for female workers are higher than for males. The uncompensated own wage elasticities are higher for big farmers who hire in labour than for small farmers and landless labourers. The effects of other variables on labour supply are also as expected. Work days increase with age but at a decreasing rate. The number of young children reduces female workdays but has no effect on male labour supply. Family size and ethnic/caste group dummy variables are also significant.

Regarding tests of model selection, there are only slight differences in the overall fit of the various labour supply models specifications and in the values of the estimated parameters. But the varying effective wage model is clearly preferred over the common wage version on the basis of standard model selection diagnostics (J tests for non-nested models and the Akaike Information Criteria). This result holds for both the male and female labour supply regressions. These model selection findings offer an independent corroboration of the efficiency differences between family and hired labour inputs in the farm production process. The evidence that the labour supply behaviour conforms to the production function estimation results supports the conclusion that there is a genuine efficiency difference between family and hired labour. The labour supply results reduce the likelihood that the estimate of $\theta < 1$ in the production function is caused by other unobserved factors not taken into account in the production function estimations.

8.5 Some Implications

Apart from the methodological issue of the appropriate specification of a farm household model that allows for heterogeneity between family and hired labour, there are several other important implications of the finding that a substantial efficiency difference exists between family and hired labour as production inputs in the *tarai* region of Nepal.

An immediate implication is with regard to the measurement of farm level efficiency. One of the celebrated "stylized facts" about agricultural production in developing countries is that small farms are cultivated more intensively (i.e. with higher levels of variable inputs used per hectare) than bigger farms. This often leads to an observed inverse relationship between farm size and productivity in terms of output per hectare. The relatively greater application of inputs on smaller farms is most pronounced in the case of labour. Per hectare labour input on small farms is consistently higher than on bigger farms over a large range of farm sizes; and this result holds whether or not average yields on small farms are higher (Berry and Cline 1979). The usual explanation for the higher labour input on smaller farms has been labour market imperfections which lead small farms to apply too much family labour to their own farms, relative to the prevailing market wages at which big farms hire in labour (Sen 1975, Carter 1992). Such "dualism" in traditional agriculture can occur if, for instance, wage rates are not bid down competitively to clear the hired labour market so that small farm family members are constrained in the amount of off-farm work they can find. The labour allocation equilibrium on the small farm is then given by the equality of the marginal product of labour with the real opportunity cost of on-farm work, at a level below the observed market wage rate. This type of market failure explains the higher labour intensity on small farms since their real cost of labour is lower than the market wage rate for hired labour applicable on the big farms.[5]

The efficiency difference between family and hired labour provides an alternative explanation for the relatively lower per hectare input of labour on the bigger farms, without relying on labour or other factor market failures. The efficiency difference means that the effective wage rate faced by big farmers who hire in labour is larger than the wage rates faced by small farmers who also work on the off-farm wage market. The small farmer equates the returns from farm cultivation to the market wage rate (w) at which he is able to work on the off-farm labour market. The big farmer faces a higher effective wage rate (w/θ) than the small farmer does in terms of equivalent units of labour because of the lower productivity of hired labour. Although family labour is applied more intensively to substitute for the less efficient hired labour, there is a rising utility cost to increased family labour application. Consequently, total per hectare labour input, measured in conventional units, is higher on small farms cultivated solely by family members than on big farms that rely on hired labour, even when wage rates are competitively determined and the hired labor market clears.

[5] This is sometimes referred to as a greater "self-exploitation" of peasant labour on small family farms (Sen 1975).

It is important to distinguish between the factor market imperfection and the effort related efficiency difference explanations because they imply different opportunity costs of labour in agriculture. For instance, under the factor market imperfection hypothesis, if a member of a small farm household migrates to an urban area, the farm output loss will be minimal. The marginal product of family labour on the small farm is hypothesized to be less than the rural market wage rate. Also, given the presumed factor market imperfections, other family members who are constrained on the off-farm labour market can possibly supply extra labour to make up for the migrating member. Under the efficiency difference hypothesis, marginal products and the real cost of leisure foregone are equated to the effective wage rates on both small and big farms, with the proviso that the effective wage rate is higher by the $1/\theta$ factor on the big farms. So the withdrawal of family labour from either the big or small farm will involve a higher opportunity cost.[6]

Theoretically, both factor market imperfections and efficiency differences could exist simultaneously. One does not rule out the other. The main focus of this thesis was not to discriminate between alternative explanations for dualism in traditional agriculture. Nevertheless, the estimation results of this study which clearly indicate labour heterogeneity, also provide strong evidence against the factor market imperfection hypothesis. Based on the estimated production function parameters in Chapter VI, the marginal product of family labour on small farms is approximately equal to the market wage rate. The labour supply regression results in Chapter VII showed the market wage accurately reflects the opportunity cost of family labour in small farms (Category 1 households, which also includes the landless). There is no direct evidence that shortage of work at the going off-farm wage forces members of small farm households to devote all their labour supply to the family farm.

---

[6] Unless, of course, the household's equilibrium labour allocation with heterogeneous labour also satisfies the "surplus labour" equilibrium in the sense defined by Sen (1966), with a flat schedule for the real cost of family labour. (The real cost of labour measures the marginal rate of indifferent substitution between consumption and leisure as represented by the VV curve in Figure 4.1B). In an initial equilibrium on a flat section of the real cost of labour schedule, the withdrawal of some family members from the agricultural labour force induces other family members to work longer hours, since there is no increasing cost to higher labour supply. Surplus labour exists in the sense that some workers can be transferred out of the farm sector without reducing labour input and farm output.

Another important implication of the labour heterogeneity finding - and one which has very significant policy relevance- is that aggregate outcomes in the agricultural sector, for instance, total production, total labour absorption, and equilibrium rural wage rates, would be sensitive to changes in the distribution of individual household endowments of land and labour. For instance, if family labour is more productive because it applies more *effort* per unit time than hired labour, then, a re-distributive land reform program which transferred land to small family-labour-operated farms from big farms relying primarily on hired labour would increase the average labour intensity of cultivation. It would increase the total absorption of labour in the agricultural sector and agricultural output without the necessity of drawing any additional resources into the agricultural sector.

8.6 Suggestions for Further Research

The main shortcoming of the research presented in this thesis is that the exact source of the efficiency difference detected between family and hired labour has not been analyzed directly. From the available data set it is not possible to discriminate between alternative explanations for the higher efficiency of family labour - i.e., differences in the intensity of effort applied, farm specific-experience of older generations, or some other unobservable heterogeneity. Data sets with more detailed information on specific characteristics and family and hired labour employed on each farm and the specific tasks they do may be able to discriminate among these alternative explanations. A key variable may be the supervisory role that family labour plays over hired labour. The MPHBS data set from Nepal Rastra Bank did not distinguish between the actual physical fieldwork and supervisory role of family labor inputs. Where this distinction is available in the data, if it can be shown that the efficiency of hired labour is improved with greater supervisory input from family members, as in (Frisvold 1994), then it would be more direct evidence for the shirking explanation for the lower efficiency of hired labour detected in our estimations.

The potential importance of the supervisory role of family labour in the presence of hired labour inputs also suggests that family labour may provide a different type of labour service – one which combines elements of "management services" with units of ordinary labour. While the later component could be easily substitutable with hired labour, the management services component is likely to be imperfectly substitutable. Further empirical research work with appropriate data sets that can distinguish these two components of family labour would be a useful extension of the results in this thesis. One viable way to deal with the problem of the essential jointness of the management services and ordinary labour units of family labour could be to estimate production functions with the labour input of the main family-based farm manager (or decision maker) as a separate labour input variable. This specification could then be used to test whether the farm manager's labour input is heterogeneous with respect to hired labour, or indeed with the labour input of other family members.

Another area for further research is to develop and estimate more elaborate specifications for the labour supply of individual household members which is consistent with labour heterogeneity. The labour supply regressions in Chapter VII were based on a simple linear model. Additional structure can be imposed on the labour supply equations to reflect an underlying utility function that allows for a more flexible non-linear response of leisure demand to wage rates. This would provide more robust results on whether the effective wage rates for family labour varies according to the hired labour market exposure of the farm households, in the presence of other possible sources of varying wage and income responses of labour supply.[7]

Other refinements to the simple labour supply specification used in this study can be made by considering a joint family labour supply model for individual members of a

[7] Stern (1986) provides a menu of alternative functional forms for labour supply functions and their underlying properties that would be a useful staring point. Only a limited number of specifications allow for a flexible wage and income response.

household, where the work days of a single person are affected by the work days put in by other household members according to some age-sex grouping (as in Newman and Gertler 1994). Another adjustment would be to test for and then incorporate constraints on the number of days for which hired work is available, given the seasonal nature of agricultural operations; and to estimate the labour supply functions with binding time constraints, where applicable, and to test for behavioural responses conforming to labour heterogeneity with this more general specification.

Another fruitful line of extension of this research would be to embed the structure and parameters of the farm household model estimated in this thesis into a larger simulation model of the agricultural economy of the *tarai* region of Nepal. The simulation model could then look specifically at the household-specific welfare effects and aggregate outcomes of alternative land reform policies that re-distributed land from big farms dependent on hired labour to family based cultivators. The general equilibrium effects, allowing for wage rates to change in response to changes in land distribution, are unclear and worthy of analysis. Of particular interest would be household-specific welfare effects. Even if the land transfers *increased* the aggregate labour absorption, there could still be a *reduction* in demand for hired labour that would lead to a reduction in the market wage rates for hired labour, and hence adversely affect landless households that were not beneficiaries of the land transfers. The direction and exact magnitude of such general equilibrium effects will depend on the precise values of the elasticities of labour demand and supply for the different classes of farm households. The research work reported in this thesis has made an initial contribution towards specifying the appropriate analytical framework and carrying out the relevant estimation work to obtain accurate estimates of these elasticities and other relevant parameters required for the simulation model.

REFERENCES

Abedin, Joynal and G. K. Bose (1988). Farm Size and Productivity Difference - A Decomposition Analysis; *Bangladesh Development Studies*, 16(3):71-79.

Abowd, J. and O. Ashenfelter (1981). Anticipated Unemployment, Temporary Layoffs, and Compensating Wage Differentials; in S.Rosen (ed.) *Studies in Labour Markets*; Chicago: Univ. of Chicago Press.

Acharya, Meena (1979). Surplus Labour in Nepalese Agriculture, *The Journal of Development and Administrative Studies* (Katmandu, Nepal), 1(2):222 243.

_________, (1987). A Study of Rural Labour Markets in Nepal, unpublished Ph.D. dissertation, University of Wisconsin Madison.

Adulavidhaya, K., Y Kuroda, L. Lau, P. Lerttamrab and P. Yotopoulos (1979). A Microeconomic Analysis of the Agriculture of Thailand, *Food Research Institute Studies*, 17: 79-86.

Ahmed, Iqbal (1981). Farm Size and Labour Use: Some Alternative Explanations, *Oxford Bulletin of Economics and Statistics*, 43:73-88.

Alderman, Harold and D. Sahn (1993). Substitution between Goods and Leisure in a Developing Country; *American Journal of Agricultural Economics*, 75(4): 875-83.

Andermann, G., H. Hockmann and G. Schmitt (1996). Historical Changes in Land-Labor Relationships in Western Europe; A paper presented at the XIII International Conference of Agricultural Economists, Sacramento, Cal., USA

Arrow, K., H. Chenery, B. Minhas and R. Solow (1961). Capital Labour Substitution and Economic Efficiency; *The Review of Economics and Statistics*, 43(3): 225-247.

Bapna, S. L., H.P Binswagner and J. Quizon (1984). Systems of Output Supply and Factor Demand Equations for Semi-Arid Tropical India; *Indian Journal of Agricultural Economics,* 39(2): 179-202.

Bardhan, Kalpana (1993). Work in South Asia: An Inter-regional Perspective; in Saraswati Raju and D. Bagchi (eds.), *Women and Work in South Asia: Regional Patterns and Perspectives*; London and New York: Routledge.

Bardhan, Pranab K. (1984). Peasant Labor Supply: A Statistical Analysis; in P. K. Bardhan, *Land Labor and Rural Poverty: Essays in Development Economics* (1984); New York: Columbia University Press.

_________, (1979). Labor Supply Functions in a Poor Agrarian Economy, *American Economic Review*, 69(1):73-83.

_________, (1973). Size, Productivity and Returns to Scale: An Analysis of Farm Level Data in Indian Agriculture; *Journal of Political Economy*, 81(6): 1370-80.

Bardhan, Pranab K and Ashok Rudra (1981). Terms and Conditions of Labour Contracts in Agriculture: Results of a Survey in West Bengal, 1979; *Oxford Bulletin of Economics and Statistics*, 43(1): 89-111.

Barnum, H. and L Squire (1979). *A Model of an Agricultural Household*; The World Bank, Washington D.C.

Barro, R. J. (1977). Unanticipated Money Growth and Unemployment in the United States; *American Economic Review*, 67:101-115.

Becker, G (1965). A Theory of the Allocation of Time; *Economic Journal* 75, 493-517

Belbase, Krishna and R. Grabowski (1985). Technical Efficiency in Nepalese Agriculture; *Journal of Developing Areas,* 19(4): 515-25.

Belbase, Krishna; R. Grabowski and O. Sanchez (1985). The Marginal Productivity of Inputs and Agricultural Production in Nepal; *Pakistan Development Review,* 24(1):51-60.

Benjamin, Dwayne (1995). Can Unobserved Land Quality Explain the Inverse Productivity Relationship?; *Journal of Development Economics*, 46(1): 51-84.

________, (1992). Household Composition, Labor Markets, and Labor Demand: Testing for Separation in Agricultural Household Models; *Econometrica*, 60(2):287-322.

Berndt, Ernst R. and L.R Christensen (1973a). The Internal Structure of Functional Relationships: Separability, Substitution and Aggregation; *Review of Economic Studies*, 40(3):403-10.

_________, (1973b). The Translog Function and the Substitution of Equipment, Structures, and Labor in U.S. Manufacturing 1929-68; *Journal of Econometrics*, 1(1): 81-113.

_________, (1974). Testing for the Existence of a Consistent Aggregate Index of Labor Inputs; *American Economic Review*, 64(3):391-404.

Berry, R.A. and W. Cline (1979). *Agrarian Structure and Productivity in Developing Countries*; Baltimore: Johns Hopkins Univ. Press.

Bhalla, Surjit and P.Roy (1988). Mis-specification in Farm Productivity Analysis: The Role of Land Quality; *Oxford Economic Papers*, 40:55-73.

Binswagner, Hans P. and M. R. Rosenzweig (1986). Behavioural and Material Determinants of Production Relations in Agriculture; *Journal of Development Studies*, 22(3):503-39.

________________, (1984). *Contractual Arrangements, Employment and Wages in Rural Labour Markets in Asia;* New Haven: Yale Univ. Press.

Binswagner, H. P. (1974). A Cost Function Approach to the Measurement of Factor Demand Elasticities of Substitution; *American Journal of Agricultural Economics*, 56:377-86.

Blackorby, Charles and R. Russell (1989). Will the Real Elasticity of Substitution Please Stand Up? (A Comparison of the Allen/Uzawa and Morishima Elasticities); *American Economic Review*, 79(4): 883-888.

Blackorby, Charles, D. Primont and R. Russell (1979). *Duality, Separability, and Functional Structure: Theory and Economic Applications*; New York: North Holland.

Blomquist, N. Soren (1983). The Effect of Income Taxation on the Labor Supply of Married Men in Sweden; *Journal of Public Economics*, 22(2):169-97.

Blundell, Richard and C. Meghir (1986). Selection Criteria for a Microeconometric Model of Labour Supply; *Journal of Applied Econometric*, 1:55-80.

Blundell, Richard and I. Walker (1982). Modelling the Joint Determination of Household Labour Supplies and commodity Demands; *The Economic Journal*, 92: 351-364.

Boisvert, R. (1982). The Translog Production Function: Its Properties, Its Several Interpretations and Estimation Problems; *Agricultural Economics Research*, No. *82-28*; Cornell University, Ithaca, N.Y.

Burtless, Gary and J. Hausman (1978). The Effect of Taxation on Labor Supply: Evaluating the Gary Negative Income Tax Experiments; *Journal of Political Economy*, 86(6):1103-30.

Caillavet, F., H. Guyomard and R. Lifran eds. (1994). *Agricultural Household Modelling and Family Economics*; Amsterdam: Elsevier.

Carter, Michael R. (1984). Identification of the Inverse Relationship between Farm Size and Productivity: An Empirical Analysis of Peasant Agricultural Production; *Oxford Economic Papers*, 36:131-46.

Chambers, R. (1986). *Applied Production Analysis: A Dual Approach*; Cambridge: Cambridge University Press.

Chayanov, A.V. (1966). *The Theory of Peasant Economy* (translation edited by Daniel Thorner, Basile Kerblay and R. E. F. Smith); Homewood, Ill., R. D. Irwin (Published for the American Economic Association).

Chiappori, Pierre A. (1988). Rational Household Labor Supply; *Econometrica*, 56(1): 63-90.

________, (1997). Introducing Household Production in Collective Models of Labor Supply; *Journal of Political Economy*, 105(1):191-209.

Cogan, John F. (1981). Fixed Costs and Labor Supply; *Econometrica*, 49(4): 945-63.

Cooke, Priscilla A. (1998). Intrahousehold Labor Allocation Responses to Environmental Good Scarcity: A Case Study from the Hills of Nepal, *Economic Development and Cultural Change*, 46(4):807-30.

Deaton, Angus (1994). *The Analysis of Household Surveys: Microeconometric Analysis for Development Policy*; (draft manuscript), Research Program in Development Studies, Princeton University,.

________, (1984). Issues in the Methodology of Multimarket Analysis of Agricultural Pricing Policies; Discussion Paper no. 116, Research Program in Development Studies, Princeton University.

Deaton, A and J. Muellbauer (1980). *Economics and Consumer Behaviour*; Cambridge: Cambridge University Press.

Diamond, Charles and T. Fayed (1998). Evidence on Substitutability of Adult and Child Labour; *Journal of Development Studies*, 34(3):62-70.

Denny, M. and M. Fuss (1977). The Use of Approximation to Test for Separability and Existence of Consistent Aggregates; *American Economic Review*, 67(3):404-418.

Diewert, W.E. (1976). Exact and Superlative Index Numbers; *Journal of Econometrics*, 4(2):115-45.

________, (1971). An Application of the Shephard Duality Theorem: A Generalized Leontief Production Function; *Journal of Political Economy*, 79(3): 481-507.

Diewert, W. E. and T. J. Wales (1989). Flexible Functional Forms and Global Curvature Conditions; *Econometrica*, 55(1):43-86.

Dixit, Avinash K. and Victor Norman (1980). *Theory of International Trade: A Dual, General Equilibrium Approach*; Welwyn, Herts: James Nisbet (Cambridge Economic Handbooks).

de Janvry, Alain, M. Fafchamps and E. Sadoulet (1991). Peasant Household Behaviour with Missing Markets: Some Paradoxes Explained; *The Economic Journal*, 101:1400-1417.

Deolalikar, Anil B. (1988). Nutrition and Labor Productivity in Agriculture: Estimates for Rural South India; *Review of Economics and Statistics*, 70(3): 406-13.

Deolalikar, Anil B and Wim P.M.Vijverberg (1987). A Test of Heterogeneity of Family and Hired Labour in Asian Agriculture; *Oxford Bulletin of Economics and Statistics*, 49(3): 291-305.

_________, (1983). The Heterogeneity of Family and Hired Labor in Agricultural Production: A Test Using District-Level Data from India; *Journal of Economic Development*, 8(2): 45-69.

Douglas, Evan J. (1989). The Simple Analytics of the Principal-Agent Incentive Contract; *Journal of Economic Education*, 20(1):39-51.

Feder, Gershon (1985). The Relation between Farm Size and Farm Productivity: The Role of Family Labor, Supervision and Credit Constraints; *Journal of Development Economics*, 18(2-3): 297-313.

Field, E.B. (1988). Free and Slave Labor in the Antebellum South: Perfect Substitutes or Different Inputs?; *The Review of Economics and Statistics*, 70(4): 654-659.

Frisvold, George B. (1994). Does Supervision Matter? Some Hypothesis Tests Using Indian Farm Level Data; *Journal of Development Economics*, 43(2):217-38.

Ghose, Ajit Kumar (1979). Farm Size and Land Productivity in Indian Agriculture: A Reappraisal; *Journal of Development Studies*, 16(1): 27-49.

Goldman, S. and H. Uzawa (1964). A Note on Separability in Demand Analysis; *Econometrica*, 32: 387-99.

Grabowski, Richard (1990). Curve Fitting versus Hypothesis Testing: A Comment; *Applied Economics*, 22(4): 427-29.

Grabowski, R. and K. Belbase (1986). An Analysis of Optimal Scale and Factor Intensity in Nepalese Agriculture: An Application of a Ray-Homothetic Production Function; *Applied Economics*, 18(10):1051-63.

Greene, William (1993). *Econometric Analysis* (2nd ed.), New York: Macmillian.

Griliches, Z. (1984). Data Problems in Econometrics; in Z. Griliches and M. Intrilligator (eds.), *Handbook of Econometrics* vol. 3; Amsterdam: North-Holland.

Gronau R (1977). Leisure, Home Production and Work - The Theory of the Allocation of Time Revisited; *Journal of Political Economy*, 85:1099-1123.

Hall, R. E. (1973). Wages, Income and Hours of Work in the U.S. Labor Force; in G. G. Cain and H.W. Watts (eds.), *Income Maintenance and Labor Supply*; Chicago: Markham.

Hamal, Krishna (1991). Crop Productivity Analysis: Nepal, 1961-1987; unpublished Ph.D. dissertation, University of Alberta (Canada).

Hammersh, Daniel (1986) Labor Demand; in Orley Ashenfelter and Richard Layard (eds.), *Handbook of Labor Economics*, vol. I; Amsterdam: Elsevier Science.

Harris, John R. and Michael P. Todaro (1970). Migration, Unemployment and Development: A Two Sector Analysis; *American Economic Review*, 60(1):126-42.

Hausman J. A. (1978). Specification Tests in Econometrics; *Econometrica,* 46:1251-1271.

________, (1981). Labour Supply, in H. J. Aaron and J. A. Pechman (eds.), *How Taxes Affect Labour Supply*; Brookings Institute; Washington DC.

Heckman, J. and Thomas Macurdy (1986). Labor Econometrics, in Z. Griliches and M. D. Intrilligator (eds.) *Handbook of Econometrics* vol. 3; Amsterdam: Elsevier.

Heckman J. J. (1979). Sample Selection Bias as a Specification Error; *Econometrica,* 47:153-161.

_______, (1980). Sample Selection Bias as a Specification Error, in James P. Smith (ed.), *Female Labor Supply: Theory and Estimation*; Princeton: Princeton Univ. Press.

Hicks, J. R. (1970). Elasticity of Substitution Again: Substitutes and Complements; *Oxford Economic Papers*, 22:289-96.

________, (1964). *The Theory of Wages* (2nd ed.); London: Macmillan.

Hirashima, S. and M. Muqtada (1986). *Hired Labour and Rural Labour Markets in Asia*; ILO, Asian Employment Programme (ARTEP), New Delhi.

HMG/N, CBS (1993). *Nepal Agricultural Census 1991/92 Report*; His Majesty's Government of Nepal, Central Bureau of Statistics, Kathmandu (various volumes).

HMG/N, MOF (1996). *Economic Survey* (of Nepal); His Majesty's Government of Nepal, Ministry of Finance, Kathmandu (annual volumes).

Ho, Teresa J. (1980). The Labour Market for Married Women in the Rural Philippine; unpublished Ph.D. dissertation, Food Research Institute, Stanford University.

Huffman, Wallace (1976). The Productive Value of Human Time in U.S. Agriculture; *American Journal of Agricultural Economics*; 58: 672-683.

Jacoby, H. (1993). Shadow Wages and Peasant Family Labour Supply: An Econometric Application to the Peruvian Sierra; *Review of Economic Studies*, 60(4): 903 21.

_______, (1991). Productivity of Men and Women and the Sexual Division of Labor in Peasant Agriculture of the Peruvian Sierra; *Journal of Development Economics*; 37(1-2): 265 87.

Jorgenson, D. and B. Fraumeni (1981). Relative Prices and Technical Change; in E. Berndt and B. Field (ed.), *Modelling and Measuring Natural Resources Substitution*; Cambdridge, Mass.: MIT Press.

Kawaguchi, Akira (1994). Testing Neoclassical and Non-neoclassical Models of Household Labour Supply; *Applied Economics*, 26(1):9-19.

Killingsworth, M. (1983). *Labour Supply*, Cambridge:Cambridge University Press.

Killingsworth, M. and J. Heckman (1986). Female Labor Supply: A Survey; in Orley Ashenfelter and Richard Layard (eds.), *Handbook of Labor Economics*, vol. I; Amsterdam: Elsevier Science.

Kuroda, Y. and P. Yotopoulos (1978). A Micro-economic analysis of Production Behaviour of the Farm Household in Japan : A Profit Function Approach; *The Economic Review* (Japan), 29:116-29.

________, (1980). A Study of Consumption Behaviour of the Farm Household in Japan: an application of the Linear Logarithmic expenditure system; *The Economic Review* (Japan), 31:1-15.

Lau, L., W Lin, P. Yotopoulos (1978). The Linear Logarithmic Expenditure System: An Application to Consumption Leisure Choice; *Econometrica*, 49:1533-51.

Laufer, Leslie A. (1985). The Substitution Between Male and Female Labor in Rural Indian Agricultural Production; *Yale Economic Growth Center Discussion Papers No. 472*; Yale University.

Leiderman, L. (1980). Macro-econometric Testing of Rational Expectations and Structural Neutrality Hypotheses for the United States; *Journal of Monetary Economics,* 6(1):69-82

Lopez, Ramon (1986). Structural Models of the Farm Household That Allow for Interdependent Utility and Profit-Maximization Decisions, in Singh, I. J., L. Squire and J. Strauss (eds.), *Agricultural Household Models: Extensions, Applications, and Policy*; Baltimore: Johns Hopkins University Press (1986):306-325.

________, (1984). Estimating Labor Supply and Production Decisions of Self-Employed Farm Producers; *European Economic Review*; 24(1): 61-82.

Maddala, G. S. (1992). *Introduction to Econometrics* (2nd edition.), New York: Macmillan.

McDonald, John and Robert Moffit (1980). The Uses of Tobit Analysis; *The Review of Economics and Statistics*, 62(2):318-321.

Mahmood, Moazam and Nadeem ul Haque (1981). Farm Size and Productivity Revisited; *Pakistan Development Review,* 20(2):151-90.

Mazumdar, Dipak (1989). Microeconomic Issues of Labor Markets in Developing Countries: Analysis and Policy Implications; EDI Seminar Paper no. 40., Economic Development Institute, World Bank, Washington D.C.

________, (1975). The Theory of Share-Cropping with Labour Market Dualism; *Economica*; 42: 261-71.

Michael, Robert and Gary Becker (1973). On the New Theory of Consumer Behavior; *Swedish Journal of Economics*; 75(4):378-96.

Moffitt, Robert (1986). The Econometrics of Piecewise-Linear Budget Constraints: A Survey and Exposition of the Maximum Likelihood Method; *Journal of Business and Economic Statistics,* 4(3):317-28.

Moghadam, Fatemet Etemad (1982). Farm Size, Management and Productivity: A Study of Four Iranian Villages; *Oxford Bulletin of Economics and Statistics,* 44(4):357-79.

Moll, Peter G. (1990). Curve Fitting vs. Hypothesis Testing with Production Functions: Comment; *Applied Economics,* 22(1):1-4.

Moschini, Giancarlo, D. Moro and R. Green (1994). Maintaining and Testing Separability in Demand Systems; *American Journal of Agricultural Economics*, 76: 61-73.

Mudhbary, P. (1988). A Demand System Analysis of Food Consumption in Nepal; unpublished Ph.D. dissertation, Michigan State University.

Murphy, Kevin and R. H. Topel (1985). Estimation and Inference in Two-step Econometric Models; *Journal of Business and Economic Statistics,* 3(4):370-379.

Mundlak, Yair (1996). Production Function Estimation: Reviving the Primal; *Econometrica*, 60(2):431:438.

___________, (1968). Elasticities of Substitution and the Theory of Derived Demand; *Review of Economic Studies*, 35: 225-236.

Nakajima, Chihiro (1949). Internal Equilibrium Theory of the Farm firm (in Japanese); *Journal of Rural Economics* (Japan), 21(3).

___________, (1986). *Subjective Equilibrium Theory of the Farm Household* (translated by R. Kada); Amsterdam: Elsevier.

Newman, John L. and Paul J. Gertler (1994). Family Productivity, Labor Supply, and Welfare in a Low Income Country; *Journal of Human Resources,* 29(4): 989-1026.

Newey, W. K. (1984). A Method of Moments Interpretation of Sequential Estimators; *Economics Letters*, 14:201-206.

NRB (Nepal Rastra Bank) (1988). *Multipurpose Household Budget Survey Report: A Study on Income Distribution, Employment and Consumption Patterns in Nepal*; Nepal Rastra Bank, Kathmandu.

NRB/ADB (1994). *Nepal Rural Credit Review Final Report* (in four volumes); Nepal Rastra Bank, Kathmandu and the Asian Development Bank, Manila.

Pagan, Adrian (1984). Econometric Issues in the Analysis of Regressions with Generated Regressors; *International Economic Review*, 25(1):221-247.

_______, (1986). Two Stage and Related Estimators and their Applications; *Review of Economic Studies,* 53:517-538.

Pencaval, J. (1986). Male Labor Supply: A Survey; in Orley Ashenfelter and Richard Layard (eds.), *Handbook of Labor Economics*, vol. I; Amsterdam: Elsevier Science.

Pitt, Mark M. and M. R. Rosenzweig (1986). Agricultural Prices, Food Consumption, and the Health and Productivity of Indonesian Farmers; in Singh, I. J., L. Squire and J. Strauss (eds.), *Agricultural Household Models: Extensions, Applications, and Policy*; Baltimore: Johns Hopkins University Press (1986); 153-182.

Quizon, J. and H. P. Binswagner (1983). Income Distribution in Agriculture: A Unified Approach; *American Journal of Agricultural Economics,* 65:526-538.

___________, (1986). Modeling the Impact of Agricultural Growth and Government Policy on Income distribution in India; *The World Bank Economic Review,* 1(1):103-148.

Rauniyar, Krishna (1985). Labour Utilization and Non Farm Labor Supply Among Rural Farm Households: A Case Study of Hill and Tarai districts; Research Paper Series no. 30, HMG/USAID/GTZ/Winrock project, Kathmandu, Nepal.

Revankar, Nagesh (1971). A Class of Variable Elasticity of Substitution Production Functions; *Econometrica*, 39(1):61-71.

Rosenzweig, M. (1980). Neoclassical Theory and the Optimizing Peasant: An Econometric Analysis of Family Labour Supply in a Developing Country; *Quarterly Journal of Economics*, 95:31-55.

____________, (1988). Labor Markets in Low-Income Countries; in H. B. Chenery, T. N. Srinivasan and Jere R. Behrman (eds.). *Handbook of Development Economics* vol. 1; Amsterdam: North Holland; 714-759.

_______, (1978). Rural Wages, Labor Supply, and Land Reform: A Theoretical and Empirical Analysis; *American Economic Review*, 68(5): 847-61.

Rosenzweig, Mark R. and Kenneth Wolpin (1985). Specific Experience, Household Structure, and Intergenerational Transfers: Farm Family Land and Labor Arrangements in Developing Countries; *Quarterly Journal of Economics,* 100(5), Supp.:961-87.

Rudra, Ashok (1992). *Political Economy of Indian Agriculture*, Calcutta: Bagchi. Publications.

Ryan, David L. and T. J. Wales (1998). A Simple Method for Imposing Local Curvature in Some Flexible Consumer Demand Systems; *Journal of Business and Economic Statistics,* 16(3):331-338.

Seidman, L. (1989). Complements and Substitutes: The Importance of Minding *p*'s and *q*'s; *Southern Economic Journal,* 56(1).

Sato, R. and T. Koizumi (1973). On the Elasticities of Substitution and Complementarity; *Oxford Economic Papers*, 25(1): 44-56.

Sen, Amartya K. (1966). Peasants and Dualism with or without Surplus Labor; *Journal of Political Economy,* 74:425-450.

__________, (1975). *Employment, Technology and Development*; Oxford: Clarendon Press.

Sharma, Kailash (1988). Farm-Household Economics and Some Empirics from the Semi-arid Tropics of India (unpublished manuscript); University of New England, Armidale; 1-11.

SHAZAM (1993). *User's Reference Manual Version 7.0*; Toronto: McGraw Hill.

Sicular, Terry (1986). Using a Farm Household Model to Analyze Labor Allocation on a Chinese Collective Farm; in Singh, I. J., L. Squire and J. Strauss (eds.) (1986), *Agricultural Household Models: Extensions, Applications, and Policy*; Baltimore: Johns Hopkins Univ. Press; 277-305.

Singh, I. J., L. Squire and J. Strauss (eds.) (1986). *Agricultural Household Models: Extensions, Applications, and Policy*; Baltimore: Johns Hopkins University Press.

Singh, I. J., L. Squire and J. Strauss (1986c1). An Overview of Agricultural Household Models; Chapter 1 in Singh, I. J., L. Squire and J. Strauss (eds.) (1986) *Agricultural Household Models: Extensions, Applications, and Policy*; Baltimore: Johns Hopkins University Press; 17-47.

Singh, I. J., L. Squire and J. Strauss (1986c2). Methodological Issues; Chapter 2 in Singh, I. J., L. Squire and J. Strauss (eds.) (1986) *Agricultural Household Models: Extensions, Applications, and Policy*; Baltimore: Johns Hopkins University Press; 48-94.

Singh, I. J., L. Squire and J. Strauss (1986b). A Survey of Agricultural Household Models: Recent findings and Policy Implications; *The World Bank Economic Review*, 1:149-179.

Skoufias, Emmanuel (1994). Using Shadow Wages to Estimate Labor Supply of Agricultural Households; *American Journal of Agricultural Economics*, 76(2):215-227.

Stern, Nicholas (1986). On the Specification of Labour Supply Functions; in R. Blundell and I. Walker (eds.), *Unemployment, Search and Labour Supply;* Cambridge: Cambridge University Press; 143-89.

Strauss, John and Duncan Thomas (1995). Human Resources: Empirical Modeling of Household and Family Decisions, in T. N. Srinivasan and Jere R. Behrman (eds.) *Handbook of Development Economics* vol. 3A:1833-2023; North Holland: Amsterdam.

Strauss, J. (1986). The Theory and Comparative Statics of Agricultural Household Models: A General Approach (Appendix to Chapter 2) in I. J. Singh, L. Squire and J. Strauss (eds.), *Agricultural Household Models: Extensions, Applications, and Policy*; Baltimore: Johns Hopkins University Press.

Squires, Dale and S. Tabor (1994). The Absorption of Labor in Indonesian Agriculture; *Developing Economies*, 32(2):167-87.

Tanaka, O. (1951). An Analyses of Economic Behaviour of the Farm-Household (in Japanese); *Journal of Rural Economics* (Japan), 22(4).

Topel, R. (1982). Inventories, Layoffs, and the Short Run Demand for Labor; *American Economic Review*, 72:769-787.

Udry, Christopher (1996). Gender, Agricultural Production, and the Theory of the Household; *Journal of Political Economy*; 104(5):1010-46.

Udry, Christopher, *et al.* (1995). Gender Differential in Farm Productivity: Implications for Household Efficiency and Agricultural Policy; *Food Policy,* 20(5):407-423.

Upadhyaya, Hari K. and Ganesh B. Thapa (1994). Modern Variety Adoption, Wage Differential, and Income Distribution in Nepal; in Cristina C. David, Keijiro Otsuka (eds.), *Modern Rice Technology and Income Distribution in Asia*; Boulder, Colorado: L. Rienner Publishers; 281-323.

Uzawa, Hirofumi (1963). Production Functions with Constant Elasticities of Substitution; *Review of Economic Studies,* 30:81-90.

Wallace, Michael B. (1989). Nepal: Food Pricing with an Open Border, in Terry Sicular (ed.), *Food Price Policy in Asia*; Ithaca: Cornell Univ. Press.

White, B. (1994). Children, Work and Child Labour: Changing Responses to the Employment of Children; *Development and Change*, 25: 849-78.

Wu, D. M. (1983). Tests of Causality, Predeterminedness and Exogeneity; *International Economic Review*, 24:547-88.

____________, (1973). Alternative Tests of Independence between Stochastic Regressors and Disturbances; *Econometrica*, 41: 733-50.

Zellner, A., J. Kmenta and J. Dreze (1966). Specification and Estimation of Cobb-Douglas Production Functions; *Econometrica*, 34:789-95.